



The present of Italian Mediterranean buffalo: precision breeding based on multi-omics data

Mayra Gómez Carpio¹, Roberta Cimmino¹, Dario Rossi¹, Gianluigi Zullo¹, Giuseppe Campanile², Gianluca Neglia², Stefano Biffani³

¹ Italian National Association of Buffalo Breeders, St Petrarca 42, 81100 Caserta, Italy

² Department of Veterinary Medicine and Animal Production Federico II University, St Federico Delpino 1, 80137 Naples, Italy

³ National Research Council (CNR), Institute of Agricultural Biology and Biotechnology (IBBA), St Via Alfonso Corti 12, 20133 Milan, Italy

ABSTRACT

Genetic evaluation in the Italian Mediterranean Buffalo (IMB) traditionally relied on the BLUP method (best linear unbiased predictor), a mixed model system incorporating both random and fixed effects simultaneously. However, recent advancements in genome sequencing technologies have opened up the opportunity to incorporate genomic information into genetic evaluations. The ssGBLUP (single-step best linear unbiased predictor) has become the method par excellence. It replaces the traditional relationship matrix with one that combines pedigree and genomic relationships, allowing for the estimation of genetic values for non-genotyped animals. The findings of this study highlight how genomic selection enhances the precision of breeding values, facilitates greater genetic advancement and reduces the generation interval, ultimately enabling a rapid return on investment.

Section: RESEARCH PAPER

Keywords: buffalo; genetic improvement; genotype; phenotype; accuracy

Citation: Mayra Gómez Carpio, Roberta Cimmino, Dario Rossi, Gianluigi Zullo, Giuseppe Campanile, Gianluca Neglia, Stefano Biffani, The present of Italian Mediterranean buffalo: precision breeding based on multi-omics data, Acta IMEKO, vol. 12, no. 4, article 33, December 2023, identifier: IMEKO-ACTA-12 (2023)-04-33

Section Editor: Leopoldo Angrisani, Università degli Studi di Napoli Federico II, Naples, Italy

Received October 11, 2023; **In final form** November 11, 2023; **Published** December 2023

Copyright: This is an open-access article distributed under the terms of the Creative Commons Attribution 3.0 License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

Funding: This work was supported by ITALIAN MINISTRY OF AGRICULTURE (MIPAAF – DISR 07) BIG” Prot. N. 0215513 11/05/2021. CUP ANASB: J29J21003720005; CUP UNINA: J69J21003020005 and by the Programma di ricerca per la Biosicurezza delle Aziende Bufaline- articolo 4 bis dell’OPCM n. 3634 del 21/12/2007.

Corresponding author: Mayra Gómez Carpio, e-mail: m.gomezcarpio@anasb.it

1. INTRODUCTION

Breeding is a method that brings about enduring and progressive enhancements in a breed's performance. In the case of buffalo, the progress of genetic selection has been hindered by challenges such as limited pedigree data, challenges in data collection, or subpar reproductive performance. As a result, the potential of buffalo has not been fully exploited. However, the Italian Mediterranean Buffalo (IMB) can be considered the only breed for this specie in the world with a reliable genetic program implemented more than 20 years ago.

Thanks to the Italian Buffalo Breeders Association (ANASB) database, it was possible to select and identify the data needed to introduce the genomic selection (GS) in the IMB, aiming to enhance the accuracy of breeding value estimation and shorten the generation interval [1].

With the advancements in molecular quantitative genetics techniques, molecular approaches for animal breeding have gained increasing importance. These methods (encompassing marker-assisted selection [2], genomic selection (GS) [3], and genome editing [4], [5]) have become increasingly crucial in the field of livestock breeding.

The constant improvements in sequencing technologies and the development of a chip specifically suited for buffalo played a huge part in reducing the GS costs [6], allowing us to genotype several subjects and perform huge progresses in buffalo breeding. Hence, a modern breeding system requires the incorporation of data from multi-omics levels [7]. High-throughput genotyping technologies facilitate precise and swift genome-wide genotyping on a large scale.

The GS method uses genome loci, specifically single nucleotide polymorphisms (SNPs), for estimating breeding

value. In contrast to conventional breeding techniques like BLUP, genomic selection breeding (ssGBLUP) has substantially enhanced the precision of breeding value estimation [8]. Among other advantages, this approach can effectively enhance traits that are difficult to measure or have low heritability. Furthermore, it not only shortens the generation interval but also reduces reproductive costs by up to 90 % [9].

Therefore, the aim of this study is to assess the efficacy of genomic models in predicting the genetic value of milk traits in the IMB.

2. METHODOLOGY

2.1. Data

Production and pedigree data for the present study were provided by the ANASB. The dataset consisted of 743'904 lactations of 276'451 buffaloes and a pedigree with 308'736 animals. As far as genotypes are concerned, genomic data of animals registered in the Italian Mediterranean Buffalo Herd Book are used. Genotyping was conducted utilizing the Axiom™ Buffalo Genotyping Array 90k microarray, specifically designed for the buffalo species. This chip allows genotyping of 90'000 points of variation across the animal's genome. After quality control and the merging of different versions of arrays, the final dataset contained approximately 47'000 SNPs corresponding to 2'250 animals.

2.2. Model

The estimate from the genetic and genomic index was performed by applying the official model used for production. Breeding values were estimated with BLUP and ssGBLUP animal models. Based on the adopted relationship matrix, two models were used: 1) the pedigree (BLUP) with the numerator relationship matrix (A); 2) the single-step genomic BLUP (ssGBLUP) in which A and the genomic relationship matrix (G) are shuffled into H, a combined relationship matrix, whose inverse (H^{-1}) was built according to Aguilar et al. [10]:

$$H^{-1} = A^{-1} + \begin{bmatrix} 0 & 0 \\ 0 & G^{-1} - A_{22}^{-1} \end{bmatrix}, \quad (1)$$

where A^{-1} and G^{-1} are the inverses of the pedigree and genomic relationship matrices, respectively; A_{22}^{-1} is the inverse of the pedigree relationship matrix for genotyped animals only. The G matrix was created according to VanRaden [11], with the following equation:

$$G = \frac{Z D Z'}{k}, \quad (2)$$

where Z is the matrix of centred gene contents, D is a diagonal matrix of SNP weights, equal to 1 in ssGBLUP, and k is the scaling parameter defined as $2 \sum p_i (1 - p_i)$ where p_i is the allele frequency of the i -th SNP. To avoid singularity, G was blended with 5 % of A_{22} [11]. Therefore, in the ssGBLUP model, SNP information is used to construct the G matrix among all the genotyped individuals.

To evaluate the result by two methods, the LR method was applied [12]. Breeding value estimation was conducted twice: first, when phenotypic data for candidate animals were completely (whole dataset), and second, when their phenotypic information was cancelled (partial dataset).

To evaluate the efficacy of the model, the following four statistics were computed:

Dispersion:

$$b_{w,p} = \frac{\text{cov}(\hat{u}_w, \hat{u}_p)}{\text{var}(\hat{u}_p)}, \quad (3)$$

Correlation:

$$\rho_{w,p} = \text{cor}(\hat{u}_w, \hat{u}_p), \quad (4)$$

Accuracy:

$$\widehat{acc} = \sqrt{\frac{\text{cov}(\hat{u}_w, \hat{u}_p)}{(1 - \bar{F}) \hat{\sigma}_u^2}}, \quad (5)$$

IncPhen [13]:

$$inc_{phen} = \frac{1}{\rho_{\hat{u}_p, \hat{u}_w}} - 1. \quad (6)$$

3. RESULTS AND DISCUSSION

The results of the precision between BLUP and ssGBLUP for the milk trait are reported in Table 1. As expected, the precision of the estimate has increased by about 3 % when using genomic information.

The crucial aspect of genetic evaluations lies in their precision, as it enables the adequate classification of animals and facilitates greater genetic progress. In the realm of genomic selection, enhancing precision is achieved through the validation of pedigrees and the meticulous implementation of quality control measures [14]. ssGBLUP enables the identification of inconsistencies within the pedigree and the adjustment of kinship relationships between genotyped and non-genotyped individuals. This adjustment plays a crucial role in reducing bias when calculating breeding values for animals without available phenotypic data [15]. On the other hand, with the ssGBLUP, inbreeding values are more precise because Mendelian segregation is estimated through observed events (SNPs) instead of relying on the expected mean probabilities incorporated in the pedigree relationship matrix [16].

Figure 1 shows how the results move in the expected direction, particularly with regard to predictive ability.

Regarding the comparison between genotyped and non-genotyped females (Table 2), the same pattern is repeated, the estimates of the genetic values are more precise and show less bias compared to the BLUP method. A very interesting result is the so-called INCPHEN (increase phenotype), which indicates when SNPs are considered in the model, the increase attributed to phenotypic data is reduced.

Usually, prioritizing the genotyping of males within a population is common practice. However, previous studies have confirmed the advantages of also including females [17], [18]. This is primarily because females play a significant role in breeding programs; they undergo selection processes and

Table 1. Accuracy statistics for milk production.

Method	Mean	Minimum	Maximum
BLUP	0.86	0.24	0.99
ssGBLUP	0.89	0.31	0.99

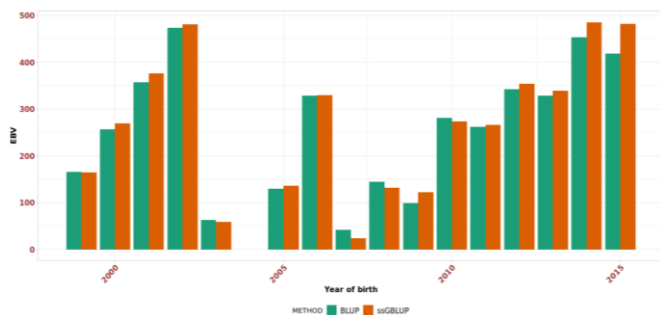


Figure 1. Histogram of EBV BLUP vs ssGBLUP.

Table 2. Validation of results with BLUP and ssGBLUP for milk production, genotyped vs. non-genotyped females.

Method	Genotyped females			
	b^a	cor^a	$incphen^a$	accuracy
BLUP	0.79	0.54	0.85	0.51
ssGBLUP	0.98	0.60	0.68	0.50
Method	Non - Genotyped females			
	b^a	cor^a	$incphen^a$	accuracy
BLUP	0.77	0.44	1.29	0.33
ssGBLUP	0.97	0.48	1.10	0.33

^a. b = Dispersion, cor = Correlation, $incphen$ = Increase phenotype

contribute valuable phenotypic data through their own performance and that of their offspring. Additionally, including females expands the genotyped population, serving as an alternative method to reduce biases and improve precision [19].

In selection programs, increasing the volume of data leads to more precise estimates. Thus, the utilization of higher-density chips tends to increase both precision and additive genetic variance [20]. While high-density chips do not capture the entirety of genetic variance, they do enhance the precision in young animals.

Regarding the economic component, several factors justify the adoption of genomic selection [15]. These include the reduction of the generation interval, increased precision in breeding value estimation, and the potential for decreased collection of phenotypic records. Among the available methods, ssGBLUP has emerged as the most approach with significant economic implications, poised to replace traditional genetic evaluation systems [14]. In addition, ssGBLUP is the most efficient alternative to simultaneously adjust information from genotyped and non-genotyped animals, especially for breeds such as IMB that currently have a low number of genotyped animals and genotyping of the whole population will still take time.

4. CONCLUSIONS

These first results, obtained by inserting the genotypes within the calculation of the genetic values, move toward the expected direction; in particular, with regard to the greater accuracy and predictive capacity.

In the era of big data, the massive genomic datasets provide breeders with abundant resources, enabling more effective exploration of the genetic foundations of complex economic traits in livestock. This development will empower breeders to swiftly and accurately select and breed livestock in alignment with their specific breeding objectives.

ACKNOWLEDGEMENT

This research was funded by ITALIAN MINISTRY OF AGRICULTURE (MIPAAF – DISR 07) - Programma di Sviluppo Rurale Nazionale 2014/2020. Caratterizzazione delle risorse genetiche animali di interesse zootecnico e salvaguardia della biodiversità. Sottomisura: 10.2 - Sostegno per la conservazione, l'uso e lo sviluppo sostenibili delle risorse genetiche in agricoltura. Project: “Bufala Mediterranea Italiana - tecnologie innovative per il miglioramento Genetico - BIG” Prot. N. 0215513 11/05/2021. CUP ANASB: J29J21003720005; CUP UNINA: J69J21003020005 and by Programma di ricerca per la Biosicurezza delle Aziende Bufaline- articolo 4 bis dell’OPCM n. 3634 del 21/12/2007.

REFERENCES

- [1] B. J. Hayes, H. A. Lewin, M. E. Goddard, The future of livestock breeding: genomic selection for efficiency, reduced emissions intensity, and adaptation, *Trends Genet.* 29(4) (2013), pp. 206-214. DOI: [10.1016/j.tig.2012.11.009](https://doi.org/10.1016/j.tig.2012.11.009)
- [2] J. C. M. Dekkers, Commercial application of marker-and gene-assisted selection in livestock: strategies and lessons, *J. Anim. Sci.* 82 (2004), pp.313-328. DOI: [10.2527/2004.8213_supplE313x](https://doi.org/10.2527/2004.8213_supplE313x)
- [3] T. H. E. Meuwissen, B. J. Hayes, M. E. Goddard, Prediction of total genetic value using genome-wide dense marker maps, *Genetics* 157(4) (2001), pp. 1819-1829. DOI: [10.1093/genetics/157.4.1819](https://doi.org/10.1093/genetics/157.4.1819)
- [4] L. Qian, M. Tang, J. Yang, Q. Wang, C. Cai, S. Jiang, H. Li, K. Jiang, P. Gao, D. Ma, Targeted mutations in myostatin by zinc-finger nucleases result in double-muscling phenotype in Meishan pigs, *Scientific reports* 5 (2015), no 14435. DOI: [10.1038/srep14435](https://doi.org/10.1038/srep14435)
- [5] J. Ruan, H. Li, K. Xu, T. Wu, J. Wei, R. Zhou, Z. Liu, Y. Mu, S. Yang, H. Ouyang, Highly efficient CRISPR/Cas9-mediated transgene knockin at the H11 locus in pigs, *Scientific Reports* 5 (2015), no 14253. DOI: [10.1038/srep14253](https://doi.org/10.1038/srep14253)
- [6] D. Iamartino, E. L. Nicolazzi, C. P. Van Tassell, J. M. Reecy, E. R. Fritz-Waters, J. E. Koltjes, S. Biffani, T. S. Sonstegard, (+ 17 more authors), Design and validation of a 90K SNP genotyping assay for the water buffalo (*Bubalus bubalis*). *PLOS ONE* 12 (2017). DOI: [10.1371/journal.pone.0185220](https://doi.org/10.1371/journal.pone.0185220)
- [7] Y.L. Yang, Z. Rong, L. Kui, Future livestock breeding: Precision breeding based on multi-omics information and population personalization, *Journal of integrative agriculture* 16(12) (2017), pp 2784-2791. DOI: [10.1016/S2095-3119\(17\)61780-5](https://doi.org/10.1016/S2095-3119(17)61780-5)
- [8] N. Ibáñez-Escriche, S. Forni, J.L. Noguera, L. Varona, Genomic information in pig breeding: Science meets industry needs, *Livest Sci* 166 (2014), pp. 94-100. DOI: [10.1016/j.livsci.2014.05.020](https://doi.org/10.1016/j.livsci.2014.05.020)
- [9] J. E. Pryce, H. D. Daetwyler, Designing dairy cattle breeding schemes under genomic selection: a review of international research, *Animal Production Science* 52(3) (2011), pp. 107-114. DOI: [10.1071/AN11098](https://doi.org/10.1071/AN11098)
- [10] I. Aguilar, I. Misztal, D. L. Johnson, A. Legarra, S. Tsuruta, T. J. Lawlor, Hot topic: A unified approach to utilize phenotypic, full pedigree, and genomic information for genetic evaluation of Holstein final score1. *J. Dairy Sci.* 93(2) (2010), pp. 743-752. DOI: [10.3168/jds.2009-2730](https://doi.org/10.3168/jds.2009-2730)
- [11] P. M. VanRaden, Efficient methods to compute genomic predictions, *J. Dairy Sci.* 91(11) (2008), pp. 4414-4423. DOI: [10.3168/jds.2007-0980](https://doi.org/10.3168/jds.2007-0980)
- [12] A. Legarra, A. Reverter, Semi-parametric estimates of population accuracy and bias of predictions of breeding values and future

- phenotypes using the LR method, *Genet Sel Evol* 50 (2018), no. 53
DOI: [10.1186/s12711-018-0426-6](https://doi.org/10.1186/s12711-018-0426-6)
- [13] M. Bermann, A. Legarra, M. K. Hollifield, Y. Masuda, D. Lourenco, I. Misztal, Validation of single-step GBLUP genomic predictions from threshold models using the linear regression method: An application in chicken mortality, *J. Anim. Breed. Genet.* 138(1) (2021), pp. 4-13.
DOI: [10.1111/jbg.12507](https://doi.org/10.1111/jbg.12507)
- [14] O. F. Christensen, P. Madsen, B. Nielsen, T. Ostersen, G. Su, Single-step methods for genomic evaluation in pigs, *Animal* 6(10) (2012), pp. 1565-1571.
DOI: [10.1017/S1751731112000742](https://doi.org/10.1017/S1751731112000742)
- [15] A. Amaya Martínez, R. Martínez Sarmiento, M. Cerón-Muñoz, Genetic evaluations in cattle using the single-step genomic best linear unbiased predictor, *Ciencia Y Tecnología Agropecuaria* 21 (2020), pp. 19-31.
DOI: [10.21930/rcta.vol21_num1_art:1548](https://doi.org/10.21930/rcta.vol21_num1_art:1548)
- [16] A. P. W. De Roos, C. Schrooten, R. F. Veerkamp, J. A. M. Van Arendonk, Effects of genomic selection on genetic improvement, inbreeding, and merit of young versus proven bulls, *J. Dairy Sci.* 94(3) (2011), pp. 1559-1567.
DOI: [10.3168/jds.2010-3354](https://doi.org/10.3168/jds.2010-3354)
- [17] A. Cesarani, S. Biffani, A. Garcia, D. Lourenco, G. Bertolini, G. Neglia, I. Misztal, N.P.P. Macciotta, Genomic investigation of milk production in Italian buffalo, *Ital. J. Anim. Sci.* 20 (2021), pp. 539-547.
DOI: [10.1080/1828051X.2021.1902404](https://doi.org/10.1080/1828051X.2021.1902404)
- [18] J.R.V. Herrera, E.B. Flores, N. Duijvesteijn, N. Moghaddar, J.H. van der Werf, Accuracy of genomic prediction for milk production traits in Philippine Dairy Buffaloes, *Frontiers in Genetics* 12 (2021), no. 682576.
DOI: [10.3389/fgene.2021.682576](https://doi.org/10.3389/fgene.2021.682576)
- [19] S. Tsuruta, I. Misztal, T.J. Lawlor, Genomic evaluations of final score for US Holsteins benefit from the inclusion of genotypes on cows. *J. Dairy Sci.* 96(5) (2013), pp. 3332-3335.
DOI: [10.3168/jds.2012-6272](https://doi.org/10.3168/jds.2012-6272)
- [20] D. Jattawa, M. A. Elzo, S. Koonawootrittriron, T. Suwanasopee, Comparison of genetic evaluations for milk yield and fat yield using a polygenic model and three genomic-polygenic models with different sets of SNP genotypes in Thai multibreed dairy cattle. *Livest. Sci.* 181 (2015), pp. 58-64.
DOI: [10.1016/j.livsci.2015.10.008](https://doi.org/10.1016/j.livsci.2015.10.008)