



## Investigating the fluctuations of residual polygenic variance on predictive ability of genomic breeding values of crossbred progeny

Somayeh Barani <sup>1</sup>, Sayed Reza Miraie-Ashtiani <sup>2✉</sup>, Ardeshir Nejati-Javaremi <sup>3</sup>, Hadi Esfandyari <sup>4</sup>, Majid Khansefid <sup>5</sup>

1. Department of Animal Science, Faculty of Agriculture, University College of Agriculture and Natural Resources, University of Tehran, Karaj, Iran. E-mail: [barani.s@ut.ac.ir](mailto:barani.s@ut.ac.ir)

2. Corresponding Author, Department of Animal Science, Faculty of Agriculture, University College of Agriculture and Natural Resources, University of Tehran, Karaj, Iran. E-mail: [ashtiani@ut.ac.ir](mailto:ashtiani@ut.ac.ir)

3. Department of Animal Science, Faculty of Agriculture, University College of Agriculture and Natural Resources, University of Tehran, Karaj, Iran. E-mail: [javaremi@ut.ac.ir](mailto:javaremi@ut.ac.ir)

4. Norwegian Beef Cattle Organizations, TYR, Hamar, Norway. E-mail: [hadi.esfandyari@tyr.no](mailto:hadi.esfandyari@tyr.no)

5. Agriculture Victoria Research, Bundoora, VIC 3083, Australia School of Applied Systems Biology, La Trobe University, Bundoora, VIC 3083, Australia, E-mail: [mkhansfid@gmail.com](mailto:mkhansfid@gmail.com)

Article Info	ABSTRACT
<b>Article type:</b> Research Article	Genomic evaluation of crossbred progeny due to their limited pedigree and lack of genotyping and performance accessibility is inevitably based on the information from purebred parental populations to increase crossbred performance. Despite the fact that single-step genomic evaluation (ssGBLUP) method can use information from both genotyped and non-genotyped animals simultaneously, leading to more accurate genomic estimated breeding values, but due to the incompatibility of the genomic and pedigree relationship matrices, it might lead to dispersion (inflation/deflation) and bias in the estimated breeding values higher than that with the BLUP method. Therefore, in this study, the prediction ability of the genomic breeding value of crossbred progeny was investigated based on the ssGBLUP method and simulation data of the purebred parents and crossbred population, considering the different scaling factor of residual polygenic variance. Based on the results of this study, the use of different ratio of residual polygenic ( $\beta$ ) variance to incorporate the genomic and pedigree relationship matrices did not considerably affect the accuracy, bias and dispersion of the predicted genomic breeding values in crossbred progeny. In addition, the effect of proportion of residual polygenic variance ( $\beta$ ) in blending the genomic and pedigree relationship matrices does not significantly affect the convergence point. Therefore, to simplify the model, the default value of $\beta$ (0.05) might be used in the inverse of the relationship matrix.
<b>Article history:</b> Received: 28 May 2023 Received in revised form: 9 September 2023 Accepted: 1 October 2023 Published online: 21 June 2024	
<b>Keywords:</b> <i>Crossbreeding,</i> <i>Single step genomic evaluation,</i> <i>Prediction ability,</i> <i>Simulation.</i>	

**Cite this article:** Barani, S., Miraie-Ashtiani, S. R., Nejati-Javaremi, A., Esfandyari, H. & Khansefid, M. (2024). Investigating the fluctuations of residual polygenic variance on predictive ability of genomic breeding values of crossbred progeny. *Iranian Journal of Animal Science*, 55 (2), 339-348. DOI: <https://doi.org/10.22059/ijas.2023.359139.653946>



© The Author(s).

DOI: <https://doi.org/10.22059/ijas.2023.359139.653946>

**Publisher:** The University of Tehran Press.

### Extended Abstract

#### Introduction

Crossbreeding, as outcrossing between breeds, is widely used in most livestock breeds to benefit from hybrid vigor effects and breed complementarity to improve performance levels of crossbred offspring. Proper crossbreeding plans can also prevent inbreeding and inbreeding depression. In fact, the main reasons to use crossbreeding plans are: 1- Using different levels of additive genetic effects between breeds, 2- Take advantage of hybrid vigor due to non-additive effects such as epistasis, dominance and over dominance and also, prevent inbreeding and inbreeding depression. However, collecting information on crossbred populations is difficult and expensive so, the breeding value of crossbreds is inevitably predicted based on the information of the

parental purebreds. The genomic evaluation of crossbred offspring in order to increase their performance, is primarily based on the information of purebred parents, while it eventually leads to the selection of superior purebreds in the parental lines (Esfandyari *et al.*, 2015). Nevertheless, A variety of factors such as genetics-environment interactions, linkage disequilibrium, non-additive genetic effects such as dominance, epistasis and imprinting, reduce the genetic correlation between these two populations, therefore the selection based on purebred parental populations will not necessarily produce crossbred offspring with higher performance. However, due to the low genetic correlation between the two populations, purebred animals do not always produce crossbred offspring with enough high performance. In general, predicting the breeding value of each population requires information from the reference population of the same population, but inevitably, genetic evaluation of crossbred uses information from mixed populations with high genetic correlation.

Additionally, Single-step Genomic Best Linear Unbiased Prediction (ssGBLUP) models can simultaneously use genotyped(G) and non-genotyped animal's information ( $A_{22}$ ) which leads to realize more accurate genomic breeding values. But then, one of the major challenges in using ssGBLUP and simultaneously using pedigree and genomic matrices of genotyped and non-genotyped animals is compromised with incompatibility between the two matrices, which can lead to bias and dispersion of genomic estimated breeding values. The compatibility of these two matrices means that the mean of diagonal and non-diagonal elements of  $G$  and  $A_{22}$  must be similar. To calculate  $H^{-1}$ ,  $G^{-1}$  is blended to a small part of a positive definite matrix, which usually includes either an Identity matrix or  $A_{22}$ , but the condition for blending  $G$  with  $A_{22}$  would be inversable and compatible (Vitezica *et al.*, 2011). One simple way to reduce the bias and decrease the predicted inflation and deflation of GEBV is to consider inbreeding coefficient in calculation of inverse of the relationship matrices and reliability, besides using the appropriate proportion of residual polygenic variance in blending of genomic and pedigree relationship matrices. So, the use of multiple populations and appropriate genomic evaluation models to predict of crossbred performance has been a subject of interest of this investigation.

### Material methods

In this study, the datasets of purebred and crossbred animals were simulated based on a sheep production system for a single trait with a heritability of 0.3. This simulation was implemented with 50K single nucleotide polymorphisms and 500 quantitative trait loci (QTLs) across the whole genome. Minor allele frequencies of markers were assumed upper than 0.05 and mutation rates of the SNPs and QTL were considered  $2.5 \times 10^{-5}$ . In this situation, the QTL allele effects are inferred from a Gamma distribution allowing for  $\alpha = 0.4$ . True breeding value (TBV) has counted the sum of the additive effects for each QTL and the observed phenotypes were computed additive effect and residual effects. The ability to predict GEBV in crossbreds was investigated using ssGBLUP method and simulated information of purebred parental and crossbred progenies. Further, we investigated the effect of proportion of residual polygenic variance in blending of genomic and pedigree relationship matrices in order to improve the prediction ability.

### Results

The results showed that using genotype of crossbred animals in combination with purebred parental population in genomic evaluation improved the prediction accuracy, when the breeding objective was to increase the crossbreed performance. In addition, considering inbreeding coefficient in calculation of the inverse of the relationship matrix and the consistency of prediction decreased the prediction bias and increased the reliability of GEBV. Furthermore, different residual polygenic variance ( $\beta$ ) values did not have a significant effect on the accuracy, bias and inflation or deflation of the prediction of GEBV in crossbred animals.

### Conclusion

The prediction accuracy, bias, and dispersion were similar across the different proportion of residual polygenic variance ( $\beta$ ) in blending the genomic and pedigree relationship matrices, therefore, for the simplicity of ssGBLUP model, it is recommended to use the default  $\beta$  (0.05).



## بررسی تغییر نسبت واریانس پلی ژنی باقیمانده بر توانایی پیش بینی ارزیابی ژنومی نتاج آمیخته

سمیه بارانی<sup>۱</sup> | سید رضا میرائی آشتیانی<sup>۲</sup> | اردشیر نجاتی جواری<sup>۳</sup> | هادی اسفندیاری<sup>۴</sup> | مجید خان سفید<sup>۵</sup>

۱. گروه علوم دامی، دانشکده کشاورزی، دانشکدگان کشاورزی و منابع طبیعی، دانشگاه تهران، کرج، ایران. رایانامه: [barani.s@ut.ac.ir](mailto:barani.s@ut.ac.ir)
۲. نویسنده مسئول، گروه علوم دامی، دانشکده کشاورزی، دانشکدگان کشاورزی و منابع طبیعی، دانشگاه تهران، کرج، ایران. رایانامه: [ashtiani@ut.ac.ir](mailto:ashtiani@ut.ac.ir)
۳. گروه علوم دامی، دانشکده کشاورزی، دانشکدگان کشاورزی و منابع طبیعی، دانشگاه تهران، کرج، ایران. رایانامه: [javaremi@ut.ac.ir](mailto:javaremi@ut.ac.ir)
۴. سازمان ملی پرورش و اصلاح نژاد گاو گوشتی (TYR)، هامار نروژ. رایانامه: [hadi.esfandyari@tyr.no](mailto:hadi.esfandyari@tyr.no)
۵. بخش تحقیقات کشاورزی، وزارت کشاورزی استرالیا در ایالت ویکتوریا، باندورا، استرالیا، گروه زیست شناسی سامانه های کاربردی، دانشگاه لاتروب، باندورا، استرالیا. رایانامه: [mkhansfid@gmail.com](mailto:mkhansfid@gmail.com)

اطلاعات مقاله	چکیده
<p><b>نوع مقاله:</b></p> <p>مقاله پژوهشی</p> <p><b>تاریخ دریافت:</b> ۱۴۰۲/۰۳/۰۷</p> <p><b>تاریخ بازنگری:</b> ۱۴۰۲/۰۶/۱۸</p> <p><b>تاریخ پذیرش:</b> ۱۴۰۲/۰۷/۰۹</p> <p><b>تاریخ انتشار:</b> ۱۴۰۳/۰۴/۰۱</p> <p><b>کلیدواژه‌ها:</b></p> <p>آمیخته گری، ارزیابی ژنومی تک مرحله‌ای، توانایی پیش بینی، شبیه سازی.</p>	<p>ارزیابی ژنومی نتاج آمیخته به دلیل محدودیت دسترسی به شجره، ژنوتیپ و عملکرد آن‌ها، معمولاً بر اساس اطلاعات جمعیت‌های خالص والدینی با هدف بهبود عملکرد نتاج آمیخته انجام می‌گیرد. به کارگیری ارزیابی ژنومی تک مرحله‌ای (ssGBLUP) علی‌رغم استفاده همزمان از اطلاعات حیوانات تعیین ژنوتیپ شده و فاقد ژنوتیپ برای افزایش توانایی پیش‌بینی، به دلیل عدم سازگاری ماتریس‌های خویشاوندی ژنومی و شجره‌ای، ممکن است موجب پراکندگی بیشتر و اریب رو به بالا ارزش اصلاحی پیش‌بینی شده نسبت به روش BLUP شود. لذا در این مطالعه توانایی پیش‌بینی ارزش اصلاحی ژنومی افراد آمیخته بر اساس اطلاعات شبیه سازی شده جمعیت‌های خالص والدینی و نتاج آمیخته با استفاده از روش ssGBLUP با در نظر گرفتن نسبت‌های مختلف واریانس پلی ژنی باقیمانده بررسی شده است. بر اساس نتیجه این پژوهش استفاده از نسبت‌های مختلف واریانس پلی ژنی باقیمانده (<math>\beta</math>) در ترکیب ماتریس خویشاوندی ژنومی و شجره‌ای، تاثیر قابل ملاحظه‌ای بر صحت، اریب، پراکندگی پیش‌بینی ارزش اصلاحی ژنومی افراد آمیخته ندارد. بعلاوه در امتزاج ماتریس‌های خویشاوندی ژنومی و خویشاوندی شجره‌ای، نسبت‌های مختلف واریانس پلی ژنی باقیمانده (<math>\beta</math>) در مدت زمان رسیدن به نقطه همگرایی اثری نداشت. بنابراین برای سادگی بیشتر مدل، استفاده از حالت پیش فرض <math>\beta</math> (معادل ۰/۰۵) در تشکیل معکوس ماتریس خویشاوندی قابل توصیه می‌باشد.</p>

**استناد:** بارانی، سمیه؛ میرائی آشتیانی، سیدرضا؛ نجاتی جواری، اردشیر؛ اسفندیاری، هادی و خان سفید، مجید (۱۴۰۳). بررسی تغییر نسبت واریانس پلی ژنی باقیمانده بر توانایی پیش بینی ارزیابی ژنومی نتاج آمیخته. *نشریه علوم دامی ایران*، ۵۵ (۲)، ۳۳۹-۳۴۸. DOI: <https://doi.org/10.22059/ijas.2023.359139.653946>

© نویسندگان.

DOI: <https://doi.org/10.22059/ijas.2023.359139.653946>

ناشر: مؤسسه انتشارات دانشگاه تهران.



## مقدمه

از آمیخته‌گری به منظور بهره‌مندی از هتروزیس و اثر تکمیل‌کنندگی نژادی در بهینه‌سازی پیشرفت ژنتیکی صفات استفاده می‌شود. در واقع دلایل اصلی به کارگیری آمیخته‌گری، (۱) استفاده از سطوح مختلف آثار ژنتیک افزایشی بین نژادی (۲) بهره‌گیری از هتروزیس ناشی از آثار غیرافزایشی مثل اپیستازی و غالبیت و فوق غالبیت است (Falconer, 1996). با این وجود به دلیل سخت و پرهزینه بودن رکوردبرداری در جمعیت آمیخته گونه‌های دامی، ارزیابی ژنومی در افراد آمیخته بیشتر بر اساس اطلاعات نژادهای خالص والدینی صورت می‌گیرد که در حقیقت منجر به انتخاب افراد خالص برتر در جمعیت خالص خواهد شد، اما استفاده از اطلاعات افراد خالص والدینی پیش‌بینی‌کننده مناسبی برای عملکرد آمیخته‌ها نخواهد بود و لزوماً این افراد نتایج آمیخته با عملکرد بالاتری تولید نخواهند کرد (Esfandyari et al., 2015). با این وجود ارزیابی بر اساس اطلاعات مشترک جمعیت‌های خالص و آمیخته به طور بالقوه پیش‌بینی دقیق‌تری برای آمیخته‌ها ارائه می‌کند (Van Grevenhof & Van der Werf, 2015). به طور کلی پیش‌بینی ارزش‌اصلاحی گروه خاصی از افراد نیازمند اطلاعات جمعیت مرجع از همان دسته از افراد است اما در افراد آمیخته دارای درصد همخونی مشابه و متفاوت ( $F_1$  در مقابل  $F_2$ ) در بهترین حالت استفاده از جمعیت مرجع جداگانه می‌تواند توصیه شود. با وجود این به دلیل سخت و پرهزینه بودن جمع‌آوری اطلاعات در جمعیت‌های آمیخته، استفاده از اطلاعات جمعیت چندنژادی و یا نژادهای خالص والدینی به عنوان یک موضوع تحقیقاتی در حال بررسی و مطالعه قرار دارد (Misztal et al., 2022).

ارزیابی ژنومی تک مرحله‌ای (ssGBLUP) با به کارگیری همزمان ماتریس‌های خویشاوندی شجره‌ای شامل  $A$  و  $A_{22}$  به عنوان روابط خویشاوندی موردانتظار و ماتریس خویشاوندی ژنومی ( $G$ ) به عنوان روابط خویشاوندی محقق شده و با استفاده از توزیع مشترک این دو منجر به برآورد صحیح‌تر ارزش اصلاحی برای حیوانات فاقد ژنوتیپ می‌شود زیرا روابط خویشاوندی حیوانات فاقد ژنوتیپ با استفاده از اطلاعات ژنوتیپی سایر افراد استنباط می‌شود (Legarra et al., 2009). لذا در این روش، معکوس ماتریس روابط خویشاوندی شجره‌ای ( $A^{-1}$ )، ژنومی ( $G^{-1}$ ) و معکوس ماتریس روابط خویشاوندی شجره‌ای برای افراد دارای ژنوتیپ ( $A_{22}^{-1}$ ) در یک ماتریس ترکیبی با عنوان  $H^{-1}$  قرار می‌گیرد. با این وجود یکی از چالش‌های اصلی به کارگیری ssGBLUP عدم سازگاری ماتریس خویشاوندی شجره‌ای و ژنومی برای افراد دارای ژنوتیپ و چگونگی وزن‌دهی به این دو ماتریس به منظور سازگار نمودن آن‌ها است. سازگاری این دو ماتریس به این معنی است که باید میانگین عناصر قطری و غیرقطری ماتریس‌های  $G$  و  $A_{22}$  مشابه باشد، در غیر این صورت نتایج پیش‌بینی GEBV اریب دار و دارای پراکندگی خواهد بود (Vitezica et al., 2011). به منظور محاسبه  $H^{-1}$ ، معکوس ماتریس خویشاوندی ژنومی ( $G^{-1}$ ) را به قسمت کوچکی از یک ماتریس معین مثبت اضافه می‌کنند که معمولاً این ماتریس معین مثبت شامل یک ماتریس یک و یا ماتریس  $A_{22}$  می‌باشد ولی شرط اضافه نمودن  $G$  به  $A_{22}$ ، معکوس پذیر بودن و سازگار بودن آن با ماتریس  $A_{22}$  و معکوس آن ( $A_{22}^{-1}$ ) خواهد بود (Vitezica et al., 2011). اضافه نمودن ماتریس معین مثبت به روابط خویشاوندی ژنومی با استفاده از تعیین ضرایب مناسب  $\alpha$  و  $\beta$  انجام می‌شود. ضریب  $\beta$  نسبتی از واریانس پلی‌ژنی باقیمانده است که ماتریس خویشاوندی ژنومی آن را در نظر نمی‌گیرد. به کارگیری ضریب  $\beta$  منجر به ایجاد یک ماتریس روابط خویشاوندی ناتکین شده که معادل برازش آثار پلی‌ژنی باقیمانده در مدل خواهد بود (Liu et al., 2014)، در این صورت ضریب  $\alpha$  نیز معادل یک منهای  $\beta$  می‌باشد. کنترل عدم‌سازگاری ماتریس خویشاوندی ژنومی و شجره‌ای با تغییر این عوامل ترکیبی منجر به افزایش همگرایی و پیش‌بینی دقیق‌تر ارزش اصلاحی ژنومی (GEBV) خواهد شد. مقادیر بهینه این عوامل برای هر صفت، مدل و یا جمعیت متفاوت می‌باشد لذا در این مطالعه توانایی پیش‌بینی

1Single Step Genomic Best Liner Unbiased Prediction

2Blending

3Non-singular

4Genomic Estimated Breeding Value

GEBV افراد آمیخته با استفاده از اطلاعات ژنوتیپ جمعیت های خالص والدینی و نتایج آمیخته با در نظر گرفتن عوامل مختلف ترکیبی بررسی شده است.

## مواد و روش ها

در این مطالعه داده های افراد خالص و آمیخته برای یک صفت با وراثت پذیری ۰/۳ مشابه سیستم پرورش گوسفند با استفاده از نرم افزار QMSim شبیه سازی شد (Sargolzaei & Schenkel, 2009). در ابتدا یک جمعیت تاریخی به اندازه ۵۰۰۰ حیوان شبیه سازی و تا نسل ۱۰۰۰ اندازه این جمعیت ثابت در نظر گرفته شد، سپس برای ایجاد عدم تعادل پیوستگی، با استفاده از گلوگاه ژنتیکی اندازه جمعیت از ۵۰۰۰ به ۲۰۰۰ در نسل ۱۵۰۰ رسید و تا نسل ۲۰۰۰ اندازه جمعیت ثابت در نظر گرفته شد. پس از آن به منظور ایجاد جمعیت های اولیه A و B و با در نظر گرفتن نسبت جنسی ۱ به ۱۰ تعداد دو نمونه به اندازه ۵۰۰ راس ماده و ۵۰ راس نر از آخرین نسل جمعیت تاریخی به صورت تصادفی نمونه گیری و سپس با در نظر گرفتن مقدار ۵۰ درصد برای نسبت تولد نر به ماده و تعداد ۵ بره به ازای هر مادر در هر نسل، جمعیت های ساخته شده تا ۲۰۰ نسل به صورت تصادفی آمیزش درون نژادی داده شدند. قابل ذکر است که تعداد نتایج به ازای هر مادر با فرض دوقلو زایی ۲۰ درصد و ۴ بار زایش در هر نسل تقریباً معادل ۵ بره در هر نسل در نظر گرفته شده است.

پس از آن به منظور توسعه این دو جمعیت و ایجاد نژادهای خالص A و B با ایجاد گلوگاه ژنتیکی دوم، تعداد ۱۰۰۰ راس ماده و ۱۰۰ راس نر از هر یک از جمعیت های اولیه A و B به صورت تصادفی نمونه گیری و تا ۱۰ نسل آمیزش درون نژادی داده شدند. در این مرحله با به کارگیری مدل بهترین پیش‌بینی کننده ناریب خطی (BLUP)، ارزش های اصلاحی بر پایه شجره برآورد و با فرض ۲۰ درصد جایگزینی برای نرها و ۵۰ درصد جایگزینی برای ماده‌ها، انتخاب افراد جایگزین بر اساس بالاترین ارزش اصلاحی پیش‌بینی شده و انتخاب حیوانات حذفی بر اساس سن بالا و تقریباً پس از یک نسل انجام شد. در نهایت برای تولید نتایج آمیخته، ۱۰۰ راس نر از هر یک از نسل های ۸، ۹ و ۱۰ نژاد A به عنوان خط پدری و ۱۰۰۰ راس ماده از هر یک از نسل های ۸، ۹ و ۱۰ نژاد B به عنوان خط مادری با همدیگر تلاقی داده شد.

به منظور شبیه سازی ژنوم، تعدادی کروموزوم اتوزوم با در نظر گرفتن پارامترهای ژنتیکی جدول ۱ شبیه سازی شده است. ارزش اصلاحی واقعی (TBV) معادل مجموع آثار افزایشی QTL ها و فنوتیپ صفت شامل TBV و آثار باقیمانده در نظر گرفته شده است. ارزیابی ژنومی با هدف بهبود عملکرد نتایج آمیخته و با استفاده از اطلاعات شجره و رکورد شبیه‌سازی شده ۱۰۲۲۰۰ راس از جمعیت های خالص والدینی و ۱۵۰۰۰ ژنوتیپ از هر سه جمعیت (A، B و C) که به نسبت مساوی در نسل های اخیر نمونه گیری شده بود انجام گرفت.

جدول ۱. پارامترها استفاده شده در شبیه سازی

مقدار	پارامتر
۰/۳	وراثت پذیری
۳۴۲۱ سانتی مورگان	طول ژنوم
۲۶	تعداد کروموزوم
۵۰۰۰۰	تراکم نشانگری
۵۰۰	تعداد جایگاه صفت کمی (QTL)
$2/5 \times 10^{-5}$	نرخ جهش QTL یا جایگاه نشانگری
۰/۰۵	فراوانی آلل نادر
$(0.14, 1/66)$ گاما	توزیع آثار افزایشی QTL

1 Best Liner Unbiased Prediction (BLUP)

2 True Breeding Value (TBV)

در این مطالعه ارزیابی ژنومی تک مرحله‌ای به منظور پیدا کردن مقادیر بهینه عوامل ترکیبی و بهترین ترکیب معکوس‌پذیر ماتریس خویشاوندی ژنومی و شجره‌ای بر اساس صحت، پراکندگی، اریب و قابلیت اطمینان ارزش اصلاحی پیش‌بینی شده<sup>۴</sup> افراد آمیخته و مدل آماری رابطه (۱) با استفاده از نرم‌افزارهای خانواده BLUPF90 (Misztal et al., 2014) بررسی شده است؛

$$y = Xb + Zu + e \quad \text{رابطه ۱}$$

که در اینجا  $y$  بردار فنوتیپ،  $b$  بردار اثر عوامل ثابت (جنس و نژاد)،  $u$  بردار آثار ژنتیک افزایشی و  $e$  بردار آثار باقیمانده است که فرض می‌شود آثار ژنتیک افزایشی و باقیمانده مستقل از هم و به ترتیب دارای توزیع نرمال با پارامترهای  $u \sim N(0, H\sigma_u^2)$  و  $e \sim N(0, I\sigma_e^2)$  هستند.  $X$  و  $Z$  به ترتیب ماتریسی است که به ترتیب حیوانات را به آثار ثابت و تصادفی مرتبط می‌کند. ماتریس  $H^{-1}$  نیز به صورت زیر محاسبه خواهد شد؛

$$H^{-1} = A^{-1} + \begin{bmatrix} 0 & 0 \\ 0 & G^{-1} - A_{22}^{-1} \end{bmatrix} \quad \text{رابطه ۲}$$

ترکیب معکوس‌پذیر ماتریس‌های خویشاوندی ژنومی و شجره‌ای با اضافه نمودن کسری از ماتریس  $A_{22}$  به ماتریس  $G$  انجام می‌شود در این صورت ماتریس خویشاوندی ژنومی معادل  $(1 - \beta)G_0 + \beta A_{22}$  خواهد بود که در اینجا  $G_0$  ماتریس خویشاوندی ژنومی قبل از اضافه نمودن ماتریس معین مثبت و  $\beta$  نسبت‌های مختلف واریانس پلی‌ژنی شامل ۰/۱، ۰/۲، ۰/۵ و ۰/۰۵ است (Pahlavan et al., 2023).

صحت ارزیابی به عنوان همبستگی رتبه‌ای ارزش اصلاحی واقعی و پیش‌بینی شده در افراد آمیخته در نظر گرفته شده است. پراکندگی پیش‌بینی ارزش اصلاحی با توجه به انحراف ضریب رگرسیون ( $b_1$ ) ارزش اصلاحی واقعی بر ارزش اصلاحی پیش‌بینی شده از یک تعیین شده است. در صورتی که مقدار ضریب رگرسیون بیش از یک باشد پیش‌بینی ارزش‌های اصلاحی دارای پراکندگی کمتر از حد است و اگر ضریب رگرسیون کمتر از یک باشد پیش‌بینی‌ها دارای پراکندگی بیش از حد خواهد بود (Tsuruta et al., 2019). اریب پیش‌بینی به عنوان انحراف میانگین ارزش اصلاحی واقعی از ارزش اصلاحی پیش‌بینی شده در نظر گرفته شده است. همچنین قابلیت اطمینان (توان دوم صحت) پیش‌بینی ارزش اصلاحی بر اساس فرمول زیر محاسبه شد (Aguilar et al., 2020).

$$rel_i = 1 - \frac{PEV_i}{(1 + F_i)\sigma_u^2} \quad \text{رابطه ۳}$$

که در اینجا  $PEV_i$  نشان‌دهنده واریانس خطای پیش‌بینی،  $\sigma_u^2$  واریانس ژنتیکی و  $F_i$  میزان همخونی فرد  $i$  ام می‌باشد.

## نتایج و بحث

در این مطالعه ابتدا مولفه‌های واریانس با استفاده از نرم‌افزار airmelf90 (Misztal et al., 2014) برآورد و قابلیت پیش‌بینی ارزیابی ژنومی برای مقادیر مختلف نسبت‌های واریانس پلی‌ژنی باقیمانده ( $\beta$ ) مورد بررسی قرار گرفت که نتایج آن در جدول ۲ نشان داده شده است. نتایج این پژوهش نشان می‌دهد که صحت ارزیابی با استفاده از همبستگی رتبه‌ای در مقادیر مختلف  $\beta$ ، مشابه و حدود ۷۵ درصد می‌باشد. به علاوه استفاده از اطلاعات ژنوتیپی افراد آمیخته منجر به افزایش صحت پیش‌بینی ژنومی می‌شود که ناشی از برآورد دقیق‌تر روابط خویشاوندی و سهم واریانس مندلی است. زیرا ارزش اصلاحی هر فرد، نصف میانگین ارزش اصلاحی والدین بعلاوه میزان انحراف مندلی آن فرد است که ماتریس روابط خویشاوندی شجره بخش واریانس مندلی آن را در نظر نمی‌گیرد. اگر چه مطالعات نشان می‌دهد زمانی که کلیه QTL‌های موثر بر صفت شناسایی شده باشد ماتریس

خویشاوندی ژنومی به جای  $A_{22}$  بهتر است با یک ماتریس یکه ترکیب شود، زیرا ممکن است که از کاهش صحت ارزیابی جلوگیری شود (Fragomeni *et al.*, 2017).

جدول ۲. مولفه های واریانس و صحت پیش‌بینی ارزش اصلاحی ژنومی با در نظر گرفتن نسبت‌های مختلف واریانس پلی‌ژنی باقیمانده (β)

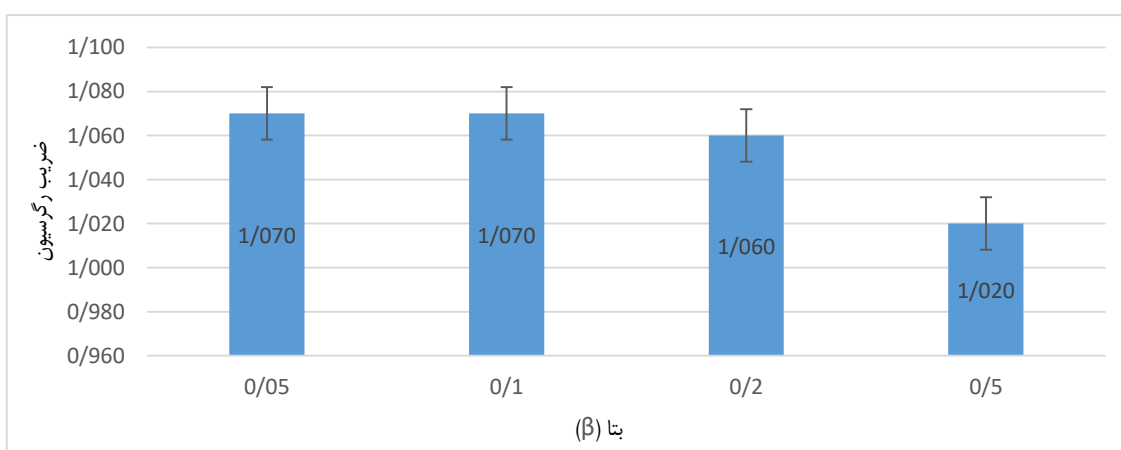
بتا (β)	قابلیت اطمینان	صحت (همبستگی رتبه‌ای)	اریب	واریانس ژنتیکی	واریانس باقیمانده
۰/۰۵	۰/۷۷۱	۰/۷۵۰	-۰/۳۷۴	۰/۱۳۰	۰/۶۹۷
۰/۱	۰/۷۷۳	۰/۷۵۳	-۰/۳۷۶	۰/۱۲۹	۰/۶۹۸
۰/۲	۰/۷۷۵	۰/۷۵۵	-۰/۳۷۶	۰/۱۲۷	۰/۶۹۸
۰/۵	۰/۷۶۵	۰/۷۴۵	-۰/۳۶۰	۰/۱۲۵	۰/۶۹۹

در این مطالعه از اطلاعات ژنوتیپ نتاج آمیخته به همراه والدین خالص به منظور پیش‌بینی ارزش اصلاحی ژنومی افراد آمیخته استفاده شده است که میزان بهبود صحت پیش‌بینی به میزان همبستگی ژنتیکی جمعیت خالص و آمیخته بستگی دارد (Esfandyari *et al.*, 2015). اگرچه در ارزیابی ژنومی گاو شیری هلشتاین، به دلیل در دسترس بودن جمعیت مرجع مناسب، فنوتیپ آمیخته‌ها در ارزیابی ژنتیکی نژاد خالص مورد استفاده قرار نمی‌گیرد زیرا این نگرانی وجود دارد که موجب کاهش صحت ارزیابی در نژاد خالص شود (VanRaden *et al.*, 2020). پیشرفت ژنتیکی ناشی از انتخاب ژنومی در گوسفند کمتر از گاو شیری بوده و به طور عمده ناشی از انتخاب برای صفاتی است که اندازه‌گیری آنها سخت و پرهزینه است (از جمله صفات مربوط به کیفیت لاشه، مقاومت به بیماری و صفات مربوط به کیفیت و میزان پشم تولیدی). زیرا بسیاری از صفات اصلاح‌نژادی در سیستم‌های مختلف پرورش گوسفند رکوردگیری نمی‌شود بنابراین عمده تلاش‌ها در توسعه انتخاب ژنومی در صنعت پرورش گوسفند به دنبال تشکیل جمعیت مرجع مناسبی است که دارای اطلاعات فنوتیپی صفات مختلف باشد (Brown *et al.*, 2018). به طور کلی استفاده از جمعیت مرجع تک نژادی در برآورد آثار نشانگری بهترین حالت در پیش‌بینی ارزش اصلاحی است، زیرا نبود پایداری در حالت عدم تعادل پیوستگی (LD) بین QTL و نشانگر در جمعیت‌های ترکیبی منجر به کاهش صحت و افزایش اریب پیش‌بینی می‌شود (Alvarenga *et al.*, 2020). از طرفی نحوه محاسبه ماتریس روابط خویشاوندی جمعیت مرجع چند نژادی به دلیل تفاوت در نسل پایه هر یک از نژادهای تشکیل‌دهنده جمعیت مرجع بر برآورد واریانس و کوواریانس ژنتیکی و همبستگی ژنتیکی بین نژادی تأثیرگذار است (Wientjes *et al.*, 2017). از طرفی هنگامی که توزیع اثرات جایگزینی آلی به فراوانی آلی هر صفت بستگی داشته باشد اثرات جایگزینی آلی صفت موردنظر در جمعیت‌های مختلف به دلیل تفاوت در فراوانی آلی متفاوت خواهد بود بنابراین برآورد روابط خویشاوندی ژنتیکی بین جمعیتی دشوار می‌باشد (Wientjes *et al.*, 2017). برای سازگار کردن ماتریس‌های خویشاوندی  $A$  و  $G$  باید ماتریس  $G$  نیز بر اساس فراوانی‌های آلی نسل پایه شجره محاسبه شود (VanRaden, 2008)، ولی مشکلی که وجود دارد عدم دسترسی به فراوانی آل‌های موردنظر می‌باشد. یکی از راه‌های ساده سازگار نمودن عناصر قطری و غیرقطری، ماتریس‌های خویشاوندی ژنومی و شجره‌ای به منظور تشکیل ماتریس  $H$ ، استفاده از ضرایب مناسب در ترکیب ماتریس  $G$  و  $A_{22}$  می‌باشد (Legarra *et al.*, 2014). تشکیل معکوس ماتریس خویشاوندی با استفاده از تغییر واریانس پلی‌ژنی نشان می‌دهد اریب پیش‌بینی ارزش اصلاحی نتاج آمیخته در کلیه سناریوها تقریباً مشابه و معادل  $-۰/۳۷$  بود (جدول ۲). منفی بودن اریب نشان‌دهنده تخمین کمتر از حد پیش‌بینی ارزش اصلاحی ژنومی است. افزایش عدم سازگاری ماتریس‌های خویشاوندی  $A$  و  $G$  در جمعیت‌های آمیخته و چند نژادی به دلیل تفاوت در فراوانی آلی و ناهمگنی جمعیت‌ها منجر به افزایش اریب و تخمین بیش از حد ارزش اصلاحی پیش‌بینی شده می‌شود (Lourenco *et al.*, 2016). زیرا میانگین عناصر قطری (میانگین همخونی) و غیر قطری (روابط خویشاوندی ژنتیکی افزایشی) در ماتریس  $A_{22}$  تحت تأثیر کامل بودن شجره، اندازه موثر جمعیت، انتخاب، رانش ژنتیکی و سیستم تلاقی‌گری قرار دارد (Falconer, 1996)، در حالی که عناصر ماتریس خویشاوندی ژنومی با استفاده از فراوانی‌های آلی برآورده می‌شود (VanRaden, 2008).

بنابراین هنگامی که ماتریس خویشاوندی ژنومی بر اساس فراوانی‌های آلی نسل‌های اخیر تخمین زده شود، ممکن است که میانگین عناصر ماتریس خویشاوندی شجره‌ای بالاتر باشد. به عبارتی میانگین GEBV بر پایه فراوانی آلی نسل‌های اخیر می‌تواند صفر باشد، حتی اگر مقدار EBV صفر نباشد. لذا ماتریس روابط خویشاوندی ژنومی بر پایه فراوانی آلی نسل‌های اخیر نمی‌تواند اثر انتخاب در نسل‌های گذشته را برآورد کند (Masuda *et al.*, 2022).

نادیده گرفتن همخونی یکی دیگر از عواملی است که باعث افزایش پراکندگی و اریب پیش‌بینی ارزش اصلاحی ژنومی با استفاده از مدل ssGBLUP می‌شود (Kluska *et al.*, 2021). از طرفی صحت مبتنی بر مدل می‌تواند به عنوان تابعی از واریانس خطای پیش‌بینی ( $PEV_i$ ) و ضریب همخونی ( $F$ ) برای هر فرد باشد، با این وجود به طور عمده میزان همخونی افراد جمعیت مورد ارزیابی در قابلیت اطمینان و محاسبه معکوس ماتریس خویشاوندی نادیده گرفته می‌شود، که این امر موجب افزایش خطای پیش‌بینی و تخمین کمتر قابلیت اطمینان ارزیابی خواهد شد (Aguilar *et al.*, 2020). عدم در نظر گرفتن همخونی در مدل ssGBLUP برای ارزش‌های اصلاحی ژنومی افراد دارای ژنوتیپ و فاقد ژنوتیپ، منجر به پیش‌بینی بیش از حد و یا کمتر از حد ارزش اصلاحی خواهد شد لذا توصیه اکید شده است که همخونی در محاسبه قابلیت اطمینان و معکوس ماتریس خویشاوندی در نظر گرفته شود (Aguilar *et al.*, 2020). در نظر گرفتن همخونی در معکوس ماتریس خویشاوندی و قابلیت اطمینان نشان می‌دهد قابلیت اطمینان برای تمامی مقادیر مختلف  $\beta$ ، مشابه و حدود ۰/۷۷ می‌باشد (جدول ۲).

پراکندگی پیش‌بینی ارزش اصلاحی ژنومی یکی از معیارهای تعیین کننده توانایی پیش‌بینی مدل است که می‌تواند ناشی از عدم سازگاری در ترکیب ماتریس خویشاوندی و ژنومی باشد (Misztal, 2017). پراکندگی بیش از حد ارزش اصلاحی هنگامی رخ می‌دهد که ضریب رگرسیون کمتر از یک است. در این صورت ممکن است کاندیدهای انتخاب در نسل‌های مختلفی قرارداد داشته باشند و یا مقادیر مختلفی از اطلاعات شجره و ژنوتیپ در دسترس باشد (Neves *et al.*, 2012). در حالی که پراکندگی کمتر از حد ارزش اصلاحی زمانی خواهد بود که ضریب رگرسیون بالاتر از یک است. نتایج میزان پراکندگی پیش‌بینی ارزش اصلاحی در مطالعه حاضر در شکل ۱ نشان می‌دهد که مقدار ۰/۵ برای  $\beta$ ، دارای کمترین پراکندگی در پیش‌بینی ارزش اصلاحی و صحت ارزیابی بود مطالعات نشان می‌دهند که استفاده از  $\beta$  کمتر از ۰/۰۰۱ منجر به افزایش اریب و عدم همگرایی می‌شود (Hollifield *et al.*, 2022). لذا بهترین دامنه  $\beta$  در ارزیابی‌های ژنومی ۰/۰۱ تا ۰/۵ گزارش شده است (Lourenco *et al.*, 2020). به علاوه استفاده از  $\beta$  بالاتر از ۰/۵ زمانی قابل توصیه است که تعداد افراد دارای ژنوتیپ کم و ماتریس خویشاوندی ژنومی نتواند واریانس خویشاوندی افزایشی را به خوبی تخمین بزند (Meyer *et al.*, 2018).



شکل ۱. مقادیر ضریب رگرسیون ارزش اصلاحی واقعی بر ارزش اصلاحی پیش‌بینی شده ژنومی (پراکندگی پیش‌بینی).



## نتیجه گیری کلی

به دلیل محدودیت در رکوردبرداری جمعیت‌های آمیخته، استفاده از مدل ارزیابی ژنومی تک مرحله‌ای فرصت مناسب برای ارزیابی ژنومی افراد آمیخته فراهم می‌کند. ولی علی‌رغم افزایش صحت پیش‌بینی، اریب و پراکندگی پیش‌بینی ارزش اصلاحی نسبت به روش BLUP بالاتر می‌باشد که به طور عمده ناشی از عدم سازگاری ماتریس روابط خویشاوندی شجره و ژنومی است. یکی از راه‌های ساده کاهش اریب و پراکندگی پیش‌بینی ارزش‌های اصلاحی استفاده از هم‌خونی در معکوس ماتریس خویشاوندی و ترکیب مناسب معکوس ماتریس  $G$  با معکوس ماتریس  $A_{22}$  خواهد بود. استفاده از مقادیر مختلف  $\beta$  تاثیر قابل ملاحظه‌ای بر صحت، اریب و تخمین بیش از حد و کمتر از حد پیش‌بینی ارزش‌های اصلاحی ژنومی افراد آمیخته نداشت لذا برای سادگی بیشتر مدل، استفاده از حالت پیش‌فرض ( $\beta$  معادل ۰/۰۵) توصیه می‌شود.

## REFERENCES

- Aguilar, I., Fernandez, E. N., Blasco, A., Ravagnolo, O., & Legarra, A. (2020). Effects of ignoring inbreeding in model-based accuracy for BLUP and SSGBLUP. *Journal of animal breeding and genetics*, 137(4), 356-364 .
- Alvarenga, A. B., Veroneze, R., Oliveira, H. R., Marques, D. B., Lopes, P. S., Silva, F. F., & Brito, L. F. (2020). Comparing alternative single-step GBLUP approaches and training population designs for genomic evaluation of crossbred animals. *Frontiers in Genetics*, 263 .
- Brown, D., Swan, A., Boerner, V., Li, L., Gurman, P., McMillan, A., Van der Werf, J., Chandler, H., Tier, B., & Banks, R. (2018). Single-step genetic evaluations in the Australian sheep industry. *Proceedings of the world congress on genetics applied to livestock production* ,
- Esfandyari, H., Sørensen, A. C., & Bijma, P. (2015). Maximizing crossbred performance through purebred genomic selection. *Genetics Selection Evolution*, 47(1), 1-16 .
- Falconer, D. S. (1996). *Introduction to quantitative genetics*. Pearson Education India .
- Fragomeni, B. O., Lourenco, D. A., Masuda, Y., Legarra, A., & Misztal, I. (2017). Incorporation of causative quantitative trait nucleotides in single-step GBLUP. *Genetics Selection Evolution*, 49(1), 1-11 .
- Hollifield, M. K., Bermann, M., Lourenco, D., & Misztal, I. (2022). Impact of blending the genomic relationship matrix with different levels of pedigree relationships or the identity matrix on genetic evaluations. *JDS communications*, 3(5), 343-347 .
- Kluska, S., Masuda, Y., Ferraz, J. B. S., Tsuruta, S., Eler, J. P., Baldi, F., & Lourenco, D. (2021). Metafounders May Reduce Bias in Composite Cattle Genomic Predictions. *Frontiers in Genetics*, 1440 .
- Legarra, A., Aguilar, I., & Misztal, I. (2009). A relationship matrix including full pedigree and genomic information. *Journal of dairy science*, 92(9), 4656-4663 .
- Legarra, A., Christensen, O. F., Aguilar, I., & Misztal, I. (2014). Single Step, a general approach for genomic selection. *Livestock Science*, 166, 54-65 .
- Liu, Z., Goddard, M., Reinhardt, F., & Reents, R. (2014). A single-step genomic model with direct estimation of marker effects. *Journal of dairy science*, 97(9), 5833-5850 .
- Lourenco, D., Legarra, A., Tsuruta, S., Masuda, Y., Aguilar, I., & Misztal, I. (2020). Single-step genomic evaluations from theory to practice: using SNP chips and sequence data in BLUPF90. *Genes*, 11(7), 790 .
- Lourenco, D., Tsuruta, S., Fragomeni, B., Chen, C., Herring, W., & Misztal, I. (2016). Crossbred evaluations in single-step genomic best linear unbiased predictor using adjusted realized

- relationship matrices. *Journal of animal science*, 94(3), 909-919 .
- Masuda, Y., VanRaden, P. M., Tsuruta, S., Lourenco, D. A., & Misztal, I. (2022). Invited review: Unknown-parent groups and metafounders in single-step genomic BLUP. *Journal of dairy science*, 105(2), 923-939 .
- Meyer, K., Tier, B., & Swan, A. (2018). Estimates of genetic trend for single-step genomic evaluations. *Genetics Selection Evolution*, 50, 1-11 .
- Misztal, I. (2017). Studies on inflation of GEBV in single-step GBLUP for type. *Interbull Bulletin* .(۵۱)
- Misztal, I., Stein, Y., & Lourenco, D. (2022). Genomic evaluation with multibreed and crossbred data. *JDS communications* .
- Misztal, I., Tsuruta, S., Lourenco, D., Aguilar, I., Legarra, A., & Vitezica, Z. (2014). Manual for BLUPF90 family of programs. *Athens: University of Georgia*, 199 .
- Neves, H. H., Carvalheiro, R., & Queiroz, S. A. (2012). A comparison of statistical methods for genomic selection in a mice population. *BMC genetics*, 13(1), 1-17 .
- Pahlavan, R., Abdollahi-Arpanahi, R., Afrazandeh, M., Nazari, B. M., & Kazemi, A. (2023). Scaling factor assessment in Single-Step GBLUP evaluations for small genotyped populations: a case study on Iranian Holstein cattle. *Livestock Science*, 105287 .
- Sargolzaei, M., & Schenkel, F. S. (2009). QMSim: a large-scale genome simulator for livestock. *Bioinformatics*, 25(5), 680-681 .
- Tsuruta, S., Lourenco, D., Masuda, Y., Misztal, I., & Lawlor, T. (2019). Controlling bias in genomic breeding values for young genotyped bulls. *Journal of dairy science*, 102(11), 9956-9970 .
- Van Grevenhof, I. E., & Van der Werf, J. H. (2015). Design of reference populations for genomic selection in crossbreeding programs. *Genetics Selection Evolution*, 47(1), 1-9 .
- VanRaden, P., Tooker, M., Chud, T., Norman, H., Megonigal Jr, J., Haagen, I., & Wiggans, G. (2020). Genomic predictions for crossbred dairy cattle. *Journal of dairy science*, 103(2), 1620-1631 .
- VanRaden, P. M. (2008). Efficient methods to compute genomic predictions. *Journal of dairy science*, 91(1), 4414-4423 .(۱)
- Vitezica, Z., Aguilar, I., Misztal, I., & Legarra, A. (2011). Bias in genomic predictions for populations under selection. *Genetics Research*, 93(5), 357-366 .
- Wientjes, Y. C., Bijma, P., Vandenplas, J., & Calus, M. P. (2017). Multi-population genomic relationships for estimating current genetic variances within and genetic correlations between populations. *Genetics*, 207(2), 503-515 .