

Comparison of the performance of deterministic formulas in predicting the accuracy of genomic evaluation in different genetic architectures

Bahareh Akbari¹, Seyed Hasan Hafezian², Mohsen Gholizadeh^{3*}

¹M.Sc. Graduate, Department of Animal Science, Faculty of Animal Science and Fisheries, Sari Agricultural Sciences and Natural Resources University, Sari, Iran.

²Professor, Department of Animal Science, Faculty of Animal Science and Fisheries, Sari Agricultural Sciences and Natural Resources University, Sari, Iran.

³Associate Professor, Department of Animal Science, Faculty of Animal Science and Fisheries, Sari Agricultural Sciences and Natural Resources University, Sari, Iran, Email: m.gholizadeh@sanru.ac.ir.

Article Info

Article type:
Research Full Paper

Article history:
Received: 02/20/2022
Revised: 04/24/2022
Accepted: 04/26/2022

Keywords:
Accuracy
Deterministic formulas
Genomic evaluation
Simulation

ABSTRACT

Background and Objectives: Identification of single nucleotide markers and different methods of genomic evaluation in the form of marker-assisted selection at the genome level has led to considerable genetic progress in the economic traits of domestic animals. The success of genomic prediction is measured by its accuracy. Deterministic formulas determine the relationship between prediction accuracy and factors affecting prediction accuracy and therefore before running into genomic selection, it is possible to design an optimal program such as the appropriate size of the reference population to achieve the optimum level of selection accuracy. The aim of the present study was to evaluate the prediction of the accuracy of deterministic formulas and compare it with the accuracy of prediction of genomic breeding values in the simulated study.

Materials and Methods: Four deterministic formulas including Daetwyler *et al* formula, Goddard formula, Goddard *et al* formula and Rabier *et al* formula were used to predict the accuracy of genomic evaluation in different genetic architectures, including different levels for heritability, reference population size, and number of independent chromosome segments. The ShinyGPAS program was used to compare and plot the accuracy of prediction. In order to compare the performance of deterministic formulas with the accuracy of predictions in the simulated population, population simulations were performed using QMSIM software. For this purpose, in genome simulation, three levels of heritability of 0.1, 0.3 and 0.5 and two levels of reference population size of 1000 and 2000 individuals were considered and estimation of genomic breeding values was performed using Bayesian method A and Bayesian B using BGLR package in R medium.

Results: In low heritability, the highest prediction accuracy was observed in the Goddard formula, which had the closest prediction accuracy (0.56) to the accuracy of genomic evaluation of simulated data estimated by the Bayes A method (0.56). With moderate heritability (0.3), Goddard (0.74) and Rabier *et al.* (0.73) had the closest and most similarity to the accuracy of the simulated data. When the population size increased from 1000 to 2000 individuals along with increasing heritability, the performance of deterministic formulas was closer to the accuracy estimated from simulation data by Bayesian methods, and the most agreement was obtained in Goddard and Rabier methods. In the lower independent chromosome segments, the highest accuracy was obtained by Rabier *et al*

(0.860). With increasing chromosomal independent segments, the highest value of accuracy was obtained by the Goddard predictive formula.

Conclusions: The results showed that deterministic formulas have a good ability to predict the accuracy of genomic evaluation and their performance is linked to the genetic architecture. The results suggest that the predictions of accuracy, in general, using Goddard and Rabier formulas are more consistent with genomic estimation accuracy in the simulated data.

Cite this article: Akbari, B., Hafezian, S.H., Gholizadeh, M. (2022). Comparison of the performance of deterministic formulas in predicting the accuracy of genomic evaluation in different genetic architectures. *Journal of Ruminant Research*, 10 (4), 39-52.



© The Author(s).

DOI: 10.22069/ejrr.2022.19965.1840

Publisher: Gorgan University of Agricultural Sciences and Natural Resources

مقایسه عملکرد معادلات قطعی در پیش‌بینی صحت ارزیابی ژنومی در ساختارهای مختلف ژنتیکی

بهاره اکبری^۱، سیدحسن حافظیان^۲، محسن قلی‌زاده^{۳*}

^۱دانش‌آموخته کارشناسی ارشد، گروه علوم دامی، دانشکده علوم دامی و شیلات، دانشگاه علوم کشاورزی و منابع طبیعی ساری

^۲استاد گروه علوم دامی، دانشکده علوم دامی و شیلات، دانشگاه علوم کشاورزی و منابع طبیعی ساری

^۳دانشیار گروه علوم دامی، دانشکده علوم دامی و شیلات، دانشگاه علوم کشاورزی و منابع طبیعی ساری، رایانامه: m.gholizadeh@sanru.ac.ir

اطلاعات مقاله	چکیده
نوع مقاله:	سابقه و هدف: شناسایی نشانگرهای تک نوکلئوتیدی و روش‌های مختلف ارزیابی ژنومی در قالب انتخاب به کمک نشانگر در سطح ژنوم منجر به پیشرفت ژنتیکی چشمگیری در صفات اقتصادی حیوانات اهلی شده است. موفقیت پیش‌بینی ژنومی بر اساس صحت آن ارزیابی می‌شود. معادله‌های قطعی، ارتباط بین صحت پیش‌بینی و عوامل مؤثر روی صحت پیش‌بینی را مشخص می‌سازند و بنابراین قبل از انجام انتخاب ژنومی، امکان طراحی برنامه بهینه مانند اندازه مناسب جمعیت مرجع برای دستیابی به سطح مطلوب صحت انتخاب را فراهم می‌نمایند. هدف از مطالعه حاضر ارزیابی پیش‌بینی صحت معادله‌های قطعی و مقایسه آن با صحت پیش‌بینی ارزش‌های اصلاحی ژنومی در مطالعه شبیه‌سازی شده بود.
تاریخ دریافت: ۱۴۰۰/۱۲/۱	
تاریخ ویرایش: ۱۴۰۱/۲/۴	
تاریخ پذیرش: ۱۴۰۱/۲/۶	
واژه‌های کلیدی:	
ارزیابی ژنومی	مواد و روش‌ها: از چهار مدل قطعی شامل معادله دتوایر و همکاران، معادله گودارد، معادله گودارد و همکاران و معادله رابیر و همکاران برای پیش‌بینی صحت ارزیابی ژنومی در ساختارهای ژنتیکی مختلف شامل طیف مختلف وراثت‌پذیری، اندازه جمعیت مرجع و تعداد قطعات مستقل کروموزومی استفاده شد. برای مقایسه و ترسیم پیش‌بینی صحت برآورد ارزش اصلاحی ژنومی از برنامه ShinyGPAS استفاده شد. به منظور مقایسه عملکرد معادلات قطعی با صحت پیش‌بینی در جمعیت شبیه‌سازی شده، شبیه‌سازی جمعیتی با استفاده از نرم‌افزار QMSIM صورت گرفت. بدین منظور در شبیه‌سازی ژنوم، سه سطح وراثت‌پذیری ۰/۱، ۰/۳ و ۰/۵ و دو سطح اندازه جمعیت مرجع ۱۰۰۰ و ۲۰۰۰ فردی در نظر گرفته شد و برآورد ارزش‌های اصلاحی ژنومی با استفاده از روش بیز A و بیز B در بسته نرم‌افزاری BGLR در محیط R انجام شد.
شبیه‌سازی صحت معادله قطعی	
	یافته‌ها: در وراثت‌پذیری‌های پایین بیشترین مقدار پیش‌بینی صحت در معادله گودارد مشاهده شد که نزدیک‌ترین پیش‌بینی صحت (۰/۵۶) را با صحت ارزیابی ژنومی داده شبیه‌سازی شده به روش بیز A (۰/۵۶) داشت. با وراثت‌پذیری متوسط (۰/۳) معادله گودارد (۰/۷۴) و رابیر و همکاران (۰/۷۳) بیشترین نزدیکی و تطابق را به صحت ارزیابی داده‌های شبیه‌سازی شده داشتند. با افزایش اندازه جمعیت از ۱۰۰۰ به ۲۰۰۰ فرد همراه با افزایش وراثت‌پذیری، عملکرد معادله‌های قطعی به صحت برآورد شده توسط روش‌های بیز نزدیک شد و بیشترین انطباق در روش گودارد و رابیر دیده شد. در قطعات مستقل کروموزومی پایین بیشترین صحت ارزش اصلاحی به دست آمده مربوط به معادله رابیر و همکاران با مقدار ۰/۸۶۰ مشاهده شد با افزایش قطعات مستقل کروموزومی، بیشترین مقدار صحت ارزش اصلاحی

به دست آمده مربوط به معادله پیش‌بینی کننده گودارد بود.

نتیجه‌گیری: نتایج تحقیق نشان داد معادله‌های قطعی از توانایی مناسبی برای پیش‌بینی صحت ارزیابی ژنومی برخوردار هستند و عملکرد آن‌ها با تغییر ساختار ژنتیکی متفاوت است. نتایج پیشنهاد می‌کنند که به‌طور کلی پیش‌بینی صحت ارزیابی ژنومی با استفاده از معادله‌های گودارد و رابیر انطباق بیشتری با صحت ارزیابی ژنومی در داده‌های شبیه‌سازی شده برخوردار هستند.

استناد: اکبری، ب.، حافظیان، س.ح.، قلی‌زاده، م. (۱۴۰۱). مقایسه عملکرد معادلات قطعی در پیش‌بینی صحت ارزیابی ژنومی در ساختارهای مختلف ژنتیکی. پژوهش در نشخوارکنندگان، ۱۰ (۴)، ۵۲-۳۹.

DOI: 10.22069/ejrr.2022.19965.1840

ناشر: دانشگاه علوم کشاورزی و منابع طبیعی گرگان

© نویسندگان.



مقدمه

مدل مختلط هندرسون امکان ارزیابی ژنتیکی را از طریق برآوردهای دقیق ارزش های اصلاحی با استفاده از مدل حیوان که در آن تمام روابط بین افراد در شجره در نظر گرفته می شود فراهم کرد. با استفاده از این مدل ها امکان برآورد همزمان اثرات ثابت و پیش-بینی اثرات تصادفی فراهم شد (Henderson, 1975).

ارزش های اصلاحی برآورد شده با استفاده از معادلات مدل های مختلط به طور وسیعی در انتخاب افراد برتر به عنوان والد ژنتیکی مورد استفاده قرار گرفتند و پیشرفت های زیادی به همراه داشتند (Silva و همکاران، 2013). با در دسترس قرار گرفتن نشانگرهای متراکم تک نوکلئوتیدی¹ که در سرتاسر ژنوم پراکنده هستند، انتخاب ژنومی معرفی شد (Meuwissen و همکاران، 2001). انتخاب ژنومی به انتخاب حیوانات بر اساس ارزش های اصلاحی ژنومی با استفاده از نشانگر اشاره می کند که در آن از تعداد بسیار زیادی از نشانگرهای ژنتیکی که کل ژنوم را پوشش می دهند به طور همزمان استفاده می شود. با استفاده از داده های شبیه سازی شده نشان داده شده است در صورتی که ارزش های اصلاحی ژنومی با استفاده از تعداد زیادی نشانگر دی ان ای پیش بینی شوند از صحت بالایی برخوردار می شوند (Meuwissen و همکاران، 2001). برخلاف انتخاب بر اساس نشانگر که در آن بخش کوچکی از واریانس ژنتیکی صفت توسط چند نشانگر بیان می شود، در انتخاب ژنومی بخش زیادی از واریانس ژنتیکی توجیه می شود (Goddard, 2009). انتخاب ژنومی در مقایسه با روش های سنتی که بر اساس رکوردهای فنوتیپی انجام می شود به دلیل کاهش فاصله نسلی، افزایش شدت انتخاب و افزایش صحت انتخاب منجر به افزایش پیشرفت ژنتیکی می شود (Solberg و همکاران، 2008). برای رسیدن به این

هدف تعداد نشانگرها باید به اندازه کافی باشد تا هر جایگاه صفت کمی حداقل با یک نشانگر در عدم تعادل لینکاژی² باشد (Toosi و همکاران، 2010). کاربرد انتخاب ژنومی در برنامه های اصلاح نژادی در حیوانات و گیاهان و پیشرفت ژنتیکی حاصل از آن به عملکرد مدل مورد استفاده در پیش بینی ارزش های اصلاحی وابسته است که تحت عنوان صحت پیش بینی همبستگی بین ارزش های اصلاحی واقعی و پیش بینی شده را بیان می کند (Morota, 2017). با فرض وجود پنل های نشانگری که به قدر کفایت متراکم هستند، صحت پیش بینی ارزش های اصلاحی ژنومی به تعداد افراد تعیین ژنوتیپ شده و تعداد افراد دارای فنوتیپ در جمعیت مرجع که در آن اثرات نشانگری برآورد می شوند، وراثت پذیری صفت و تعداد قطعات مستقل کروموزومی، تعداد نشانگرها، توزیع اثرات ژنی، توزیع فراوانی آلی و تعداد جایگاه های صفت کمی³ بستگی دارد (Daetwyler و همکاران، 2008؛ de los Campos و همکاران، 2013؛ Goddard, 2009). نقش این عوامل بر صحت پیش بینی خود تحت تأثیر انتخاب مدل پیش بینی قرار دارد (van den Berg و همکاران، 2019). برای مثال تعداد جایگاه صفت کمی صحت پیش بینی روش های بیز را تحت تأثیر قرار می دهد ولی روی بهترین پیش-بینی ناریب خطی ژنومی تأثیر ندارد (Daetwyler و همکاران، 2010). یک جمعیت مرجع با طراحی صحیح اثر بسیار زیادی در صحت پیش بینی ارزش های اصلاحی افراد جوان و بدون رکورد فنوتیپی، از طریق نقش مهم آن در برآورد اثرات نشانگری دارد. همچنین عواملی همچون قابلیت اعتماد اطلاعات فنوتیپی افراد، تعداد افراد، روابط ژنتیکی درون جمعیت مرجع و روابط خویشاوندی بین افراد جمعیت مرجع با افراد

2- Linkage disequilibrium (LD)

3- Quantitative trait loci

1- Single nucleotide polymorphism (SNP)

در این مطالعه عملکرد معادلات پیش‌بینی کننده صحت ارزیابی ژنومی در ساختارهای مختلف ژنومی و جمعیتی بررسی شد. بدین منظور از ۵ معادله پیش‌بینی کننده زیر استفاده شد و قابلیت هر کدام در ساختارهای مختلف ژنتیکی و جمعیتی مورد بررسی قرار گرفت. معادله دت‌وایلر و همکاران: این معادله مطابق زیر تابعی از وراثت‌پذیری و اندازه جمعیت مرجع و قطعات مستقل کروموزومی است (Daetwyler و همکاران، ۲۰۰۸؛ Daetwyler و همکاران، ۲۰۱۰):

$$r = \sqrt{\frac{Nh^2}{Nh^2 + M_e}}$$

که N تعداد افراد در جمعیت مرجع، h^2 وراثت‌پذیری صفت و M_e تعداد قطعات مستقل کروموزومی هستند. معادله گودارد: این معادله با تصادفی در نظر گرفتن نشانگرها و فرض عدم تعادل لینکاژی کامل بین نشانگر و جایگاه صفت کمی، معادله زیر را ارائه شد (Goddard، ۲۰۰۹):

$$r = \sqrt{1 - \frac{\gamma}{2N\sqrt{\alpha}} \ln\left(\frac{1 + \alpha + 2\sqrt{\alpha}}{1 + \alpha - 2\sqrt{\alpha}}\right)}$$

در این معادله، γ برابر است با $(M_e / (h^2 \ln(2n_e)))$ و α برابر است با $1 + 2(m_e / nh^2 \ln(2n_e))$ و N_e اندازه مؤثر جمعیت است.

معادله گودارد و همکاران: این معادله، عدم تعادل لینکاژی ناقص بین نشانگرها جایگاه صفت کمی و این واقعیت که تعداد نشانگرها محدود است در نظر می‌گیرد (Goddard و همکاران، ۲۰۱۱):

$$r = \sqrt{b \frac{Nb h^2 / M_e}{1 + Nb h^2 / M_e}}$$

در این معادله $b = M / M + M_e$ نسبت واریانس ژنتیکی بیان‌شده توسط نشانگرها و M تعداد نشانگرهاست. در صورتی که مقدار b برابر یک باشد این معادله تعیین‌کننده با معادله دت‌وایلر و همکاران برابر می‌شود.

جمعیت تأیید نقش کلیدی بر صحت پیش‌بینی‌ها در جمعیت مرجع دارد (Silva و همکاران، ۲۰۱۳).

معادلات قطعی^۱ ارتباط بین صحت پیش‌بینی و عوامل مؤثر روی صحت پیش‌بینی را مشخص می‌سازند و از این راه اجازه انتخاب طرح بهینه برای جمعیت‌های مرجع (اندازه جمعیت مرجع) برای دستیابی به سطح مطلوب صحت انتخاب کاندیداهای انتخاب را فراهم می‌نمایند (Morota، ۲۰۱۷). این پیش‌بینی کننده‌های قطعی اگر صحتی که ارائه می‌نمایند. با داده‌ها و مشاهدات در جمعیت‌های واقعی موافق و نزدیک باشند دارای ارزش زیادی در طراحی برنامه‌های اصلاح نژادی بر اساس انتخاب ژنومی دارند (Morota، ۲۰۱۷). در یک مطالعه شبیه‌سازی شده و در قیاس با جمعیت واقعی، عملکرد پیش‌بینی‌های قطعی در صحت ارزش‌های اصلاحی ژنومی پیش‌بینی شده برای کاندیداهای انتخاب بدون فنوتیپ بر اساس وراثت‌پذیری صفت، تعداد فنوتیپ و ژنوتیپ در جمعیت مرجع، اندازه جمعیت مؤثر و طول ژنوم بررسی شد (Hayes و همکاران، ۲۰۰۹). نتیجه نشان داد که پیش‌بینی کننده‌های قطعی راهنمای مفیدی برای ایجاد اندازه جمعیت‌های مرجع هستند که برای پیش‌بینی ارزش‌های اصلاحی ژنومی با صحت مطلوب در کاندیداهای انتخاب ضروری می‌باشند. هدف از این مطالعه ارزیابی صحت پیش‌بینی معادله‌های قطعی در ساختارهای مختلف ژنتیکی بود. همچنین صحت پیش‌بینی ارزش‌های اصلاحی ژنومی برآورد شده توسط معادلات پیش‌بینی کننده، با صحت پیش‌بینی ارزش‌های اصلاحی ژنومی برآورد شده توسط مطالعه شبیه‌سازی شده مورد مقایسه قرار گرفت.

مواد و روش‌ها

1-Deterministic formula

کروموزوم انجام گرفت که طول هرکدام ۱۰۰ سانتی مورگان با ۱۰۰۰ نشانگر در نظر گرفته شد. تعداد جایگاه صفت کمی روی هر کروموزوم ۱۰۰ در نظر گرفته شد که به طور تصادفی روی کروموزوم قرار داشتند و توزیع اثر آنها نرمال بود. صحت پیش بینی ارزش های ژنومی برای این شبیه سازی در وراثت پذیری به مقدار ۰/۱، ۰/۳ و ۰/۵ مورد آزمون قرار گرفت. برآورد ارزش های اصلاحی ژنومی در شبیه سازی ژنومی با استفاده از روش بیز A و بیز B انجام شد. شبیه سازی جمعیتی با استفاده از نرم افزار QMSIM (Schenkel و Sargolzaei، ۲۰۰۹) و برآورد ارزش های اصلاحی ژنومی با استفاده از روش بیز A و بیز B در بسته نرم افزاری BGLR (de los Campos و همکاران، ۲۰۱۳) در محیط R انجام شد.

نتایج و بحث

نتایج حاصل از مقایسه معادلات پیش بینی کننده صحت ارزیابی ژنومی و عملکرد دو روش بیز A و B در وراثت پذیری مختلف، در شکل ۱ ارائه شده است. در تمام معادلات با افزایش وراثت پذیری صحت برآورد ارزش های اصلاحی افزایش یافت. با افزایش وراثت پذیری و در وراثت پذیری های بالا معادله رایبیر نسبت به سایر معادله های قطعی، صحت پیش بینی ارزش اصلاحی بالاتری را نشان داد و در وراثت پذیری های پایین بیشترین مقدار صحت پیش بینی در معادله گودارد (۰/۷۱۵) مشاهده شد. همچنین همان طور که ملاحظه می شود روش آماری بیز B نسبت به بیز A صحت پیش بینی ارزش اصلاحی بالاتری را نشان داد که اختلاف بین آنها در وراثت پذیری پایین (۰/۱) چشمگیر بود و معادله گودارد نزدیک ترین عملکرد (۰/۵۶) را به صحت ارزیابی ژنومی داده شبیه سازی شده به روش بیز A (۰/۵۶) داشت. با وراثت پذیری متوسط (۰/۳) معادله

معادله رایبیر و همکاران: این معادله با رها کردن فرض $\sigma_g^2 = 1$ و $\sigma_e^2 = 1$ که در معادله دات وایلر و همکاران لحاظ می شود، به دست آمد. این معادله می تواند برای هر مقدار σ_g^2 و σ_e^2 اعمال شود (Rabier و همکاران، ۲۰۱۶):

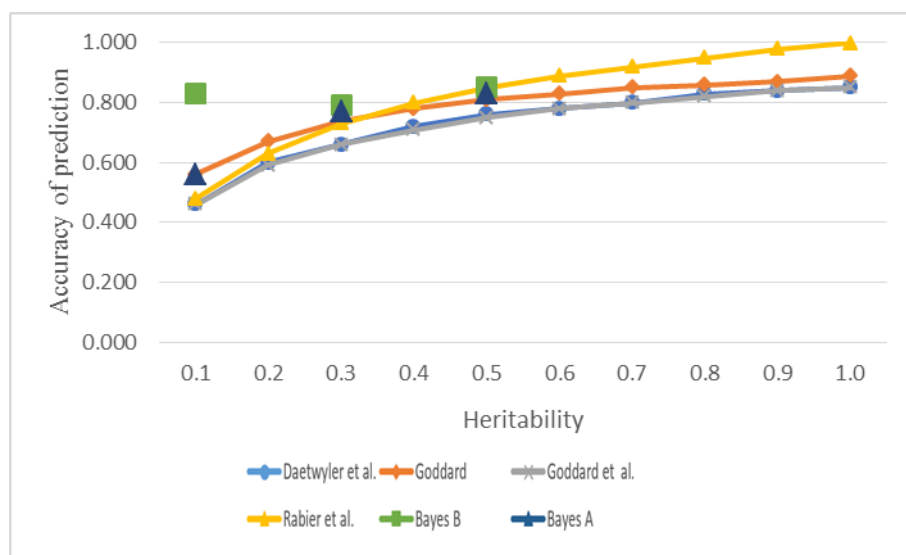
$$r = \sqrt{\frac{h^2/(1-h^2)}{m_e/n + h^2(1-h^2)}}$$

برای ترسیم و مقایسه صحت برآورد ارزش اصلاحی ژنومی معادلات پیش بینی از برنامه ShinyGPAS (Morota، ۲۰۱۷) استفاده شد. در این برنامه، ارتباط بین صحت پیش بینی و عوامل مؤثر روی صحت پیش بینی در قالب مجموعه ای از پارامترهای موجود مانند وراثت پذیری، اندازه جمعیت مرجع، تعداد قطعات مستقل کروموزومی و اندازه جمعیت مؤثر بررسی شد و نتایج برای فهم آسان، به صورت شکل ارائه شد.

شبیه سازی جمعیت و ژنوم: صحت پیش بینی ارزش های اصلاحی ژنومی برآورد شده توسط معادلات پیش بینی کننده، با صحت پیش بینی ارزش های اصلاحی ژنومی پیش بینی شده توسط مطالعه شبیه سازی شده مورد مقایسه قرار گرفت. شبیه سازی جمعیت با جمعیت پایه با تعداد ۵۰ نر و ۵۰ ماده شروع شد که با آمیزش تصادفی به مدت ۵۰ نسل ادامه داشت و جمعیت تاریخی را برای ایجاد عدم تعادل لینکاژی بر اساس تعادل جهش رانش ژنتیکی تشکیل داد. در نسل ۵۱ اندازه جمعیت به ۱۰۰۰ و ۲۰۰۰ فرد افزایش یافت که به عنوان جمعیت مرجع مورد استفاده قرار گرفت. جمعیت مرجع دارای اطلاعات فنوتیپی و ژنوتیپی بودند. نسل ۵۲ نسل جمعیت کاندیدای انتخاب بود که در آن افراد دارای اطلاعات نشانگری بودند و برای آنها ارزش های اصلاحی ژنومی پیش بینی شد. شبیه سازی ژنوم با ۴

گودارد (۰/۸۱) و رابیر و همکاران (۰/۸۵) بیشترین نزدیکی و تطابق را به صحت ارزیابی داده‌های شبیه‌سازی شده توسط دو روش بیز A (۰/۸۳) و بیز B (۰/۸۵) داشتند

گودارد (۰/۷۴) و رابیر و همکاران (۰/۷۳) بیشترین نزدیکی و تطابق را به صحت ارزیابی داده‌های شبیه‌سازی شده توسط دو روش بیز A (۰/۷۷) و بیز B (۰/۷۹) داشتند. در وراثت‌پذیری بالا (۰/۵) نیز معادله



شکل ۱: صحت ارزش‌های اصلاحی ژنومی در معادله‌های قطعی مختلف برای مقادیر مختلف وراثت‌پذیری و صحت ارزش‌های اصلاحی ژنومی برآورد شده با روش‌های بیزی در داده شبیه‌سازی شده توسط بیز B و بیز A در وراثت‌پذیری‌های مختلف

Figure 1. Accuracy of genomic breeding values from the different deterministic formulas in different values for heritability, and accuracy of the genomic breeding value estimated using Bayesian methods in simulation data

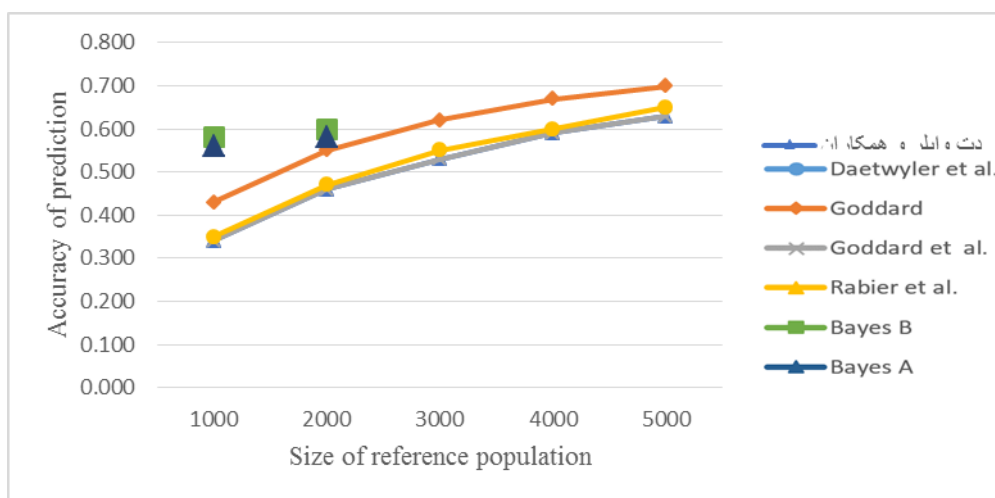
وراثت‌پذیری قابلیت اعتماد بین ارزش اصلاحی و ارزش فنوتیپی است و واضح است که هر چه بیشتر باشد فنوتیپ از قابلیت اعتماد بیشتر برخوردار شده و صحت پیش‌بینی افزایش می‌یابد. در وراثت‌پذیری پایین سهم واریانس ژنتیکی از واریانس فنوتیپی پایین است و بخش اعظم واریانس فنوتیپی تحت کنترل عوامل محیطی است.

پیش‌بینی صحت معادله‌های قطعی حاضر در این مقاله با صحت پیش‌بینی داده‌های شبیه‌سازی شده‌ای که اثرات جایگاه صفت کمی دارای توزیع نرمال بودند مقایسه شدند. معادله دتوایلر و همکاران در اکثر سناریوهای مورد بررسی کمترین پیش‌بینی صحت و بیشترین فاصله را با صحت پیش‌بینی‌های

نتایج تحقیق حاضر به‌طور کلی نتایج نشان داد که عملکرد و نزدیکی معادلات پیش‌بینی کننده صحت ارزیابی ژنومی به ساختار ژنتیکی صفت بستگی دارد. در مطالعه‌ای که با هدف بررسی عملکرد دو معادله قطعی دتوایلر و همکاران و گودارد انجام شد، نشان داده شد که هر دو معادله توافق نسبت خوبی با صحت‌های ارزیابی ژنومی در جمعیت‌های واقعی دارند و با تعداد مشابه قطعات مستقل کروموزومی و تعداد مشابه فنوتیپ‌ها در وراثت‌پذیری‌های پایین تا متوسط، معادله دتوایلر صحت کمتری را پیش‌بینی می‌کند (Hayes و همکاران، ۲۰۰۹). در تحقیق حاضر صحت ارزیابی ژنومی همه مدل‌های قطعی پیش‌بینی کننده با افزایش وراثت‌پذیری افزایش یافت.

(Goddard, 2009). در واقع تفاوت معادله گودارد و دتویلر و همکاران در این است که صحت برآورد اثر برای جایگاه صفت کمی فراوانی الل متوسط بیشتر است چرا که واریانس ژنتیکی بیشتری را نسبت به جایگاه صفت کمی با فراوانی آلی حد نهایی توضیح می‌دهد ولی، معادله دتویلر و همکاران صحت برآورد اثرات جایگاه صفت کمی را صرف نظر از فراوانی آن‌ها برابر فرض می‌کند. برای صفاتی با جایگاه صفت کمی بزرگ اثر شناخته شده مانند درصد چربی، معادله قطعی گودارد صحت پیش بینی ژنومی را دست پایین پیش بینی می‌کنند. برای اکثر صفات، اما، روش‌های پیش بینی شده ارزش اصلاحی ژنومی که توزیع اثرات جایگاه صفت کمی را نرمال فرض می‌کنند به خوبی روش‌هایی عمل می‌کنند که این توزیع را نمایی فرض می‌کنند (Hayes و همکاران، 2009).

شبه‌سازی شده داشت. این معادله اثرات نشانگرهای ژنتیکی را ثابت فرض می‌کند و استقلال جایگاه صفت کمی را جزئی از فرض دانسته و با فرض $\sigma_g^2 + \sigma_e^2 = 1$ و $\sigma_e^2 = 1$ با روش رگرسیون بر یک جایگاه در یک زمان کار می‌کند (Daetwyler و همکاران، 2008؛ Daetwyler و همکاران، 2010). همچنین معادله دتویلر دقت یکسان در برآوردهای اندازه اثر جایگاه صفت کمی در فراوانی‌های مختلف ارائه می‌کند و عدم تعادل لینکاژی کاملی بین نشانگر و جایگاه صفت کمی فرض می‌کند که این شرایط ممکن است حضور نداشته باشند. در معادله گودارد فرض بر این است که اثرات جایگاه‌های صفت کمی از توزیع نرمال نمونه‌گیری شدند و جایگاه صفت کمی با فراوانی الل کمیاب بالا، اندازه اثر دقیق‌تری نسبت به جایگاه صفت کمی با فراوانی الل کمیاب پایین دارد



شکل ۲- صحت ارزش‌های اصلاحی ژنومی در معادله‌های قطعی مختلف برای مقادیر مختلف اندازه جمعیت رفرنس در وراثت‌پذیری ۰/۱ و صحت ارزش‌های اصلاحی ژنومی برآورد شده با روش‌های بیزی در داده شبه‌سازی شده توسط بیز B و بیز A در اندازه جمعیت مؤثر ۱۰۰۰ و ۲۰۰۰.

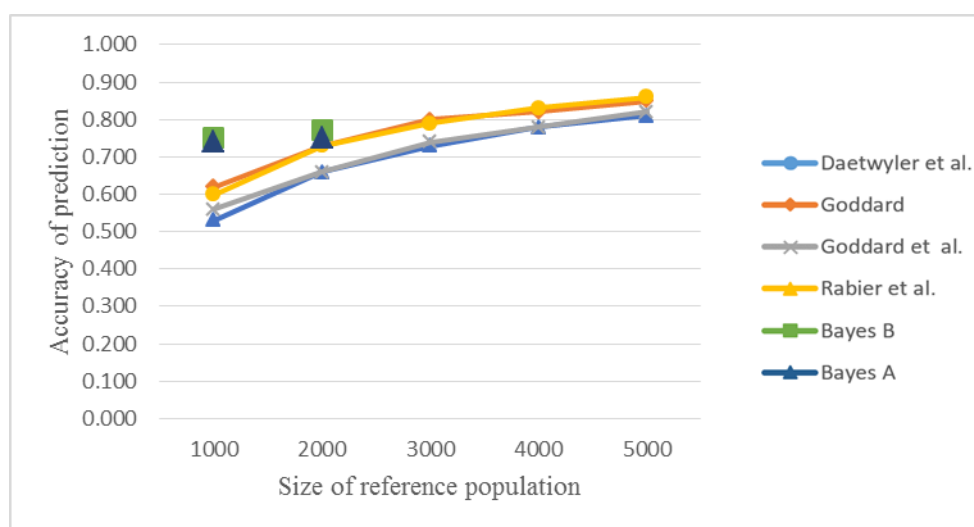
Figure 2. Accuracy of genomic breeding values from the different deterministic formulas in different values for the size of reference population and heritability of 0.1, and accuracy of the genomic breeding value estimated using Bayesian methods in simulation data with the reference population size of 1000 and 2000

مختلف نشان می‌دهد، با افزایش اندازه‌ی جمعیت مرجع از ۱۰۰۰ به ۵۰۰۰ فرد به‌طور کلی صحت ارزش‌های اصلاحی ژنومی در تمام معادلات افزایش

در این مطالعه نقش اندازه جمعیت در عملکرد معادلات پیش‌بینی کننده در وراثت‌پذیری‌های مختلف مورد بررسی قرار گرفت. نتایج حاصل از معادلات

متوسط (۰/۳) نیز با افزایش با افزایش اندازه جمعیت از ۱۰۰۰ به ۲۰۰۰ فرد، عملکرد معادلات قطعی در پیش‌بینی صحت با مقدار برآورد شده توسط روش‌های بیز نزدیک شد و بیشترین انطباق در روش گودارد و رابیر و همکاران دیده شد. در این وراثت‌پذیری عملکرد معادله‌های دتوایلر و گودارد ۲۰۱۱ به هم نزدیک بود (شکل ۳).

یافت. در وراثت‌پذیری پایین (۰/۱) و با اندازه جمعیت مرجع ۱۰۰۰ فردی مقادیر پیش‌بینی مدل‌های قطعی از صحت روش‌های بیز در جمعیت شبیه‌سازی شده کمتر بود و با افزایش اندازه جمعیت از ۱۰۰۰ به ۲۰۰۰ فرد، عملکرد معادلات قطعی در پیش‌بینی صحت به مقدار برآورد شده توسط روش‌های بیزی نزدیک شد و بیشترین انطباق در روش گودارد دیده شد (شکل ۲) در وراثت‌پذیری



شکل ۳- صحت ارزش‌های اصلاحی ژنومی در معادله‌های قطعی مختلف برای مقادیر مختلف اندازه جمعیت رفرنس در وراثت‌پذیری ۰/۳ و صحت ارزش‌های اصلاحی ژنومی برآورد شده با روش‌های بیزی در داده شبیه‌سازی شده توسط بیز B و بیز A در اندازه جمعیت مؤثر ۱۰۰۰ و ۲۰۰۰

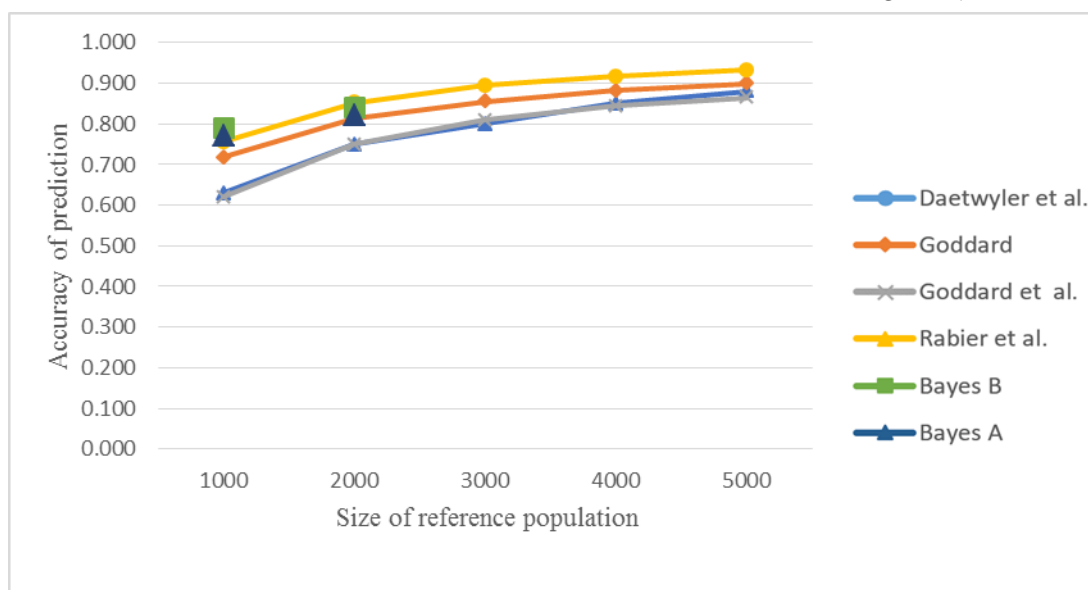
Figure 3. Accuracy of genomic breeding values from the different deterministic formulas in different heritability for the size of reference population and heritability of 0.3, and accuracy of the genomic breeding value estimated using Bayesian methods in simulation data with the reference population size of 1000 and 2000

در مطالعه حاضر مقدار اندازه جمعیت مؤثر برای معادلات پیش‌بینی کننده صحت و جمعیت شبیه‌سازی شده قابل کنترل بود ولی نزدیکی نتایج حاصله به جمعیت واقعی به مقدار این آماره در جمعیت واقعی و نزدیکی آن به مقدار محاسباتی در معادلات پیش‌بینی کننده بستگی دارد (Hayes و همکاران، ۲۰۰۹). گزارش شده است که با افزایش اندازه جمعیت مؤثر در سطوح مختلف وراثت‌پذیری (کم، متوسط و بالا)، صحت پیش‌بینی ارزش اصلاحی

با افزایش مقدار وراثت‌پذیری و در وراثت‌پذیری بالا (۰/۵) انطباق عملکرد معادله‌های پیش‌بینی شده در اندازه جمعیت مرجع ۱۰۰۰ نفری نیز دیده شد. در اندازه جمعیت ۱۰۰۰ فردی معادله پیش‌بینی کننده رابیر و در اندازه جمعیت مرجع ۲۰۰۰ فردی معادله پیش‌بینی کننده گودارد بیشترین انطباق را با صحت برآورد شده روش‌های بیز در مطالعه شبیه‌سازی نشان دادند (شکل ۴).

ژنتیکی که توسط نشانگرها تبیین می‌شود و صحت برآورد اثرات نشانگری در جمعیت رفرنس است و با افزایش اندازه جمعیت صحت برآورد اثرات نشانگری افزایش می‌یابد (Goddard, 2017). مطالعات شبیه‌سازی گذشته نیز نشان می‌دهد که با افزایش اندازه جمعیت رفرنس صحت پیش‌بینی افزایش می‌یابد (Combs و Bernardo, 2013; VanRaden, 2008).

کاهش می‌یابد به طوری که کمترین کاهش در وراثت‌پذیری‌های پایین و بیشترین کاهش در وراثت‌پذیری بالا اتفاق می‌افتد (Weber و همکاران, 2012). اندازه و ترکیب جمعیت رفرنس از عوامل مهم تأثیرگذار روی صحت پیش‌بینی هستند (van den Berg و همکاران, 2019). در مطالعه حاضر با افزایش اندازه جمعیت مرجع، صحت پیش‌بینی‌ها افزایش یافت. صحت پیش‌بینی تحت تأثیر مقدار واریانس



شکل ۴- صحت ارزش‌های اصلاحی ژنومی در معادله‌های قطعی مختلف برای مقادیر مختلف اندازه جمعیت رفرنس در وراثت‌پذیری ۰/۵ و صحت ارزش‌های اصلاحی ژنومی برآورد شده با روش‌های بیزی در داده شبیه‌سازی شده توسط بیز B و بیز A در اندازه جمعیت مؤثر ۱۰۰۰ و ۲۰۰۰

Figure 4. Accuracy of genomic breeding values from the different deterministic formulas in different values for the size of reference population and heritability of 0.5, and accuracy of the genomic breeding value estimated using Bayesian methods in simulation data with the reference population size of 1000 and 2000

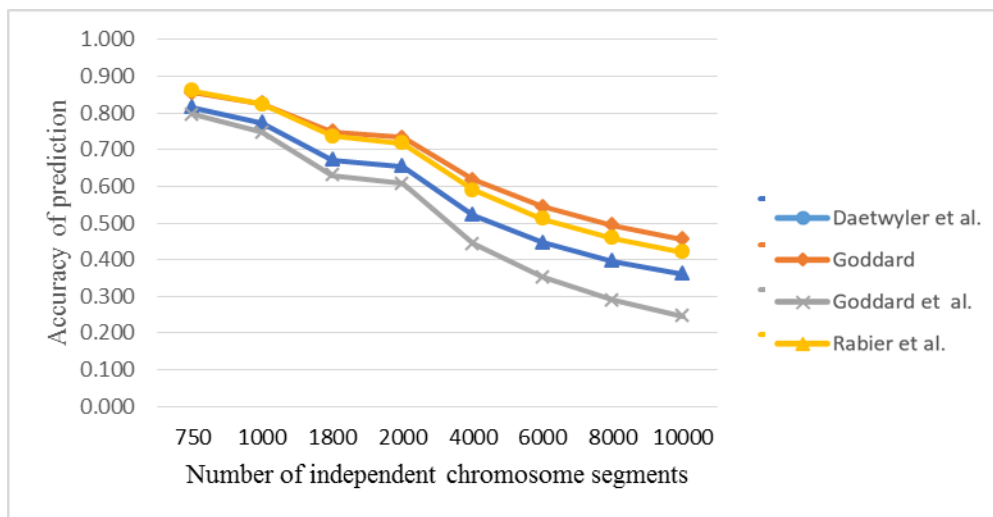
اندازه و روابط خویشاوندی کنترل داشت و اندازه جمعیت رفرنس از مهم‌ترین عوامل کنترل‌کننده صحت پیش‌بینی هستند (de los Campos و همکاران, 2013). همچنین روش آماری مورد استفاده نیز بر صحت پیش‌بینی نقش دارد. در این مطالعه برای برآورد صحت جمعیت شبیه‌سازی شده از دو روش بیز A و بیز B استفاده شد و روش بیز B در سناریوهای مورد مطالعه عملکرد بهتری داشت؛ اما تفاوت بین مدل‌ها همیشه با داده‌های واقعی تأیید نشده‌اند.

همچنین نشان داده شده است که با کاهش اندازه جمعیت مرجع تا ۶۸ درصد، مقدار صحت از ۵۳ درصد به ۳۵ درصد کاهش می‌یابد و افزایش صحت به صورت تابعی از اندازه جمعیت مرجع به صورت خطی و تا حداکثر اندازه جمعیت مرجع امکان‌پذیر است (VanRaden, 2008). اگرچه برخی از عوامل مؤثر روی صحت پیش‌بینی مانند ساختار ژنتیکی و سطح بالای عدم تعادل لینکاژی را نمی‌توان کنترل کرد ولی می‌توان روی طراحی جمعیت رفرنس از قبیل

شکل ۵ ارائه شده است. در تمام معادلات با افزایش میزان قطعات مستقل کروموزومی صحت برآورد ارزش های اصلاحی کاهش یافت (شکل ۵). در قطعات مستقل کروموزومی پایین ($Me = 750$) بیشترین صحت ارزش اصلاحی به دست آمده مربوط به معادله رابیر و همکاران با مقدار 0.86 مشاهده شد با افزایش Me بیشترین مقدار صحت ارزش اصلاحی به دست آمده مربوط به معادله پیش بینی کننده گودارد بود.

آنالیزهای تجربی فقط تفاوت های کوچکی را بین روش های مختلف نشان داده اند که در آن ها سودمندی اندکی برای روش های مبتنی بر انتخاب و جمع شونددگی مانند بیز B برای صفات با جایگاه صفت کمی بزرگ اثر نشان داده اند و روش هایی مانند بیز A برای همه صفات سودمندی یکسانی داشته اند (de los Campos و همکاران، ۲۰۱۳).

نتایج حاصل از مقایسه معادلات پیش بینی کننده صحت ارزیابی ژنومی قطعات مستقل کروموزومی در



شکل ۵- صحت ارزش های اصلاحی ژنومی در معادله های قطعی مختلف برای مقادیر مختلف ژنومی قطعات مستقل کروموزومی

Figure 5- Accuracy of genomic breeding values from the different deterministic formulas in different numbers of independent chromosome segments

است که برای برآورد دقیق قطعات مستقل کروموزومی، حداقل ۵۰ فرد باید تعیین ژنوتیپ شوند. در مطالعه حاضر تعداد قطعات مستقل کروموزومی در سناریوهایی که اثرات عوامل دیگر مطالعه شد برابر ۷۵۰ در نظر گرفته شد و همین مقدار برای شروع مطالعه همین اثر بر صحت پیش بینی در نظر گرفته شد. در مطالعاتی که روی جمعیت های گاو شیری گزارش شده است، تعداد قطعات مستقل کروموزومی به طور معنی داری به نژاد و روش برآورد این آماره بستگی داشت زمانی که در این مطالعات از Ne برای

مطالعه حاضر نشان داد که با افزایش قطعات مستقل کروموزومی صحت پیش بینی در همه مدل ها کاهش می یابد که با نتایج مطالعه دیگر روی این موضوع (VanRaden, ۲۰۰۸) مطابقت دارد. تعداد قطعات مستقل کروموزومی به صورت مستقیم به اندازه جمعیت مؤثر مرتبط است (Goddard, ۲۰۱۷؛ Hayes و همکاران، ۲۰۰۹). هر چه اندازه جمعیت مؤثر کمتر باشد مقدار عدم تعادل لینکاژی بیشتر بوده و تعداد قطعات مستقل کروموزومی کاهش و صحت پیش بینی افزایش می یابد (Marjanovic, ۲۰۲۱). نشان داده شده

به خصوص در وراثت پذیری متوسط به بالا نزدیک تر بودند. با افزایش اندازه جمعیت از ۱۰۰۰ به ۲۰۰۰ فرد، عملکرد معادلات قطعی در پیش بینی صحت با مقدار برآورد شده توسط روش های بیز نزدیک شد و بیشترین انطباق در روش گودارد و Rabier و همکاران (۲۰۱۶) دیده شد. با افزایش قطعات مستقل کروموزومی صحت پیش بینی در همه مدل ها کاهش یافت. با توجه به برآورد تعداد قطعات مستقل از اندازه جمعیت مؤثر آزمون عملکرد این مدل ها در جمعیت های واقعی پیشنهاد می شود.

برآورد Me استفاده شده است، Me درون جمعیت ها از ۸۰۰ تا ۸۰۰۰ متفاوت بوده است (Brard و Ricard, ۲۰۱۵). از آنجایی که Me به اندازه جمعیت مؤثر ارتباط دارد، قابل انتظار است که جمعیت با تنوع ژنتیکی کمتر تعداد قطعات مستقل کروموزومی کمتری داشته باشد.

نتیجه گیری کلی

معادله های قطعی مورد استفاده در این تحقیق عملکردهای مختلفی برای صحت پیش بینی ارزش های اصلاحی نشان دادند. معادله های پیش بینی کننده گودارد و رابیر و همکاران به صحت های پیش بینی ارزش های اصلاحی جمعیت شبیه سازی شده

منابع

- Brard, S. and Ricard, A. 2015. Is the use of formulae a reliable way to predict the accuracy of genomic selection. *Journal of Animal Breeding and Genetics*, 132(3): 207–217.
- Combs, E. and Bernardo, R. 2013. Accuracy of genomewide selection for different traits with constant population size, heritability, and number of markers. *The Plant Genome*, 6(1): 11.0030.
- Daetwyler, H.D., Pong-Wong, R., Villanueva, B. and Woolliams, J.A. 2010. The impact of genetic architecture on genome-wide evaluation methods. *Genetics*, 185(3): 1021–1031.
- Daetwyler, H.D., Villanuev, B. and Woolliams, J.A. 2008. Accuracy of predicting the genetic risk of disease using a genome-wide approach. *PLOS ONE*, 3(10): e3395.
- de los Campos, G., Hickey, J.M. Pong-Wong, R., Daetwyler, H.D. and Calus, M.P.L. 2013. Whole-genome regression and prediction methods applied to plant and animal breeding. *Genetics*, 193(2): 327–345.
- Goddard, M. 2009. Genomic selection: prediction of accuracy and maximisation of long term response. *Genetica*, 136(2): 245–257.
- Goddard, M. E. 2017. Can we make genomic selection 100% accurate. *Journal of Animal Breeding and Genetics*, 134: 287–288.
- Goddard, M.E., Hayes, B.J. and Meuwissen, T.H.E. 2011. Using the genomic relationship matrix to predict the accuracy of genomic selection. *Journal of Animal Breeding and Genetics*, 128(6): 409–421.
- Hayes, B.J., Daetwyler, H.D., Bowman, P., Moser, G., Tier, B., Crump, R., Khatkar, M., Raadsma, H.W. and Goddard, M.E. 2009. Accuracy of genomic selection: comparing theory and results. *Proceeding of Association for Advancement of Animal Breeding and Genetics*, 18(18): 34–37.
- Hayes, B.J., Bowman, P.J., Chamberlain, A.C., Verbyla, K. and Goddard, M.E. 2009. Accuracy of genomic breeding values in multi-breed dairy cattle populations. *Genetics Selection Evolution*, 41(1): 1–9.
- Henderson, C.R. 1975. Best linear unbiased estimation and prediction under a selection model. *Biometrics*, 423–447.
- Marjanovic, J. and Calus, M.P.L. 2021. Factors affecting accuracy of estimated effective number of chromosome segments for numerically small breeds. *Journal of Animal Breeding and Genetics*, 138(2): 151–160.

- Meuwissen, T.H.E., Hayes, B.J. and Goddard, M.E. 2001. Prediction of total genetic value using genome-wide dense marker maps. *Genetics*, 157(4): 1819–1829.
- Morota, G. 2017. ShinyGPAS: interactive genomic prediction accuracy simulator based on deterministic formulas. *Genetics Selection Evolution*, 49(1): 91.
- Pérez, P. and de Los Campos, G. 2014. Genome-wide regression and prediction with the BGLR statistical package. *Genetics*, 198(2): 483–495.
- Rabier, C.E., Barre, P., Asp, T. and Charmet G. Mangin B. 2016. On the accuracy of genomic selection. *PloS One*, 11(6): e0156086.
- Sargolzaei, M. and Schenkel, F.S. 2009. QMSim: a large-scale genome simulator for livestock. *Bioinformatics*, 25(5): 680–681.
- Silva, G.G.Z., Dutilh, B.E., Matthews, T.D., Elkins, K., Schmieder, R., Dinsdale, E.A. and Edwards, R. A. 2013. Combining de novo and reference-guided assembly with scaffold_builder. *Source Code for Biology and Medicine*, 8(1): 1–5.
- Solberg, T.R., Sonesson, A.K., Woolliams, J.A. and Meuwissen, T.H.E. 2008. Genomic selection using different marker types and densities. *Journal of Animal Science*, 86(10): 2447–2454.
- Toosi, A. Fernando, R.L. and Dekkers, J.C.M. 2010. Genomic selection in admixed and crossbred populations. *Journal of Animal Science*, 88(1): 32–46.
- van den Berg, I., Meuwissen, T.H.E., MacLeod, I.M. and Goddard, M.E. 2019. Predicting the effect of reference population on the accuracy of within, across, and multibreed genomic prediction. *Journal of Dairy Science*, 102(4): 3155–3174.
- VanRaden, P.M. 2008. Efficient methods to compute genomic predictions. *Journal of Dairy Science*, 91(11): 4414–4423.
- VanRaden, P.M. and Sullivan, P.G. 2010. International genomic evaluation methods for dairy cattle. *Genetics Selection Evolution*, 42(1): 1-7.
- Weber, K.L., Thallman, R.M., Keele, J.W. Snelling, W.M., Bennett, G.L., Smith, T.P.L., McDanel, T.G., Allan, M.F., Van Eenennaam, A.L. and Kuehn, L.A. 2012. Accuracy of genomic breeding values in multibreed beef cattle populations derived from deregressed breeding values and phenotypes. *Journal of Animal Science*, 90(12): 4177–4190.