

The effect of using different levels of noncausal markers along with causal loci and different genetic architectures on the accuracy of predicting genomic values in livestock

**Masoumeh Saleh¹, Ayoub Farhadi^{2*}, Rostam Abdollahi-Arpanahi³,
Mohsen Gholizadeh², Mahmoud Amiri-Roodbar⁴**

¹ PhD student, Department of Animal Science, Faculty of Animal Science and Fisheries, Sari Agriculture Sciences & Natural Resources University, Iran

² (Corresponding author) Associate Professor, Department of Animal Science, Faculty of Animal Science and Fisheries, Sari Agriculture Sciences & Natural Resources University, Science, Email: ayoub_farhadi@ymail.com

³ Department of Animal Sciences, University of Florida

⁴ Research Assistant (Ph.D in Animal Genetics & Breeding), Department of Animal Science, Safiabad-Dezful Agricultural and Natural Resources Research and Education Center, Agricultural Research, Education & Extension Organization (AREEO), Dezful, Iran,

Article Info

Article type:
Research Full Paper

ABSTRACT

Background and Objectives: The most important factors affecting the accuracy of genomic prediction include the number of markers, statistical method, minor allele frequency (MAF), heritability, and genetic architecture, that each of their effects differs under the specific conditions. The aim of this study was to investigate the effect of adding noncausal SNPs to the statistical model on the accuracy of breeding values by considering different population factors and different genetic architectures.

Materials and Methods: In the present study, 10 chromosomes with a length of 100 cM were simulated and on each chromosome 1000 SNP markers of two alleles were considered which were randomly distributed along the genome. In this study, phenotypic variance was assumed 1.0 and heritability was considered as 0.1, 0.2 and 0.5 according to traits with low, medium and high heritability respectively. The population numbers for this simulation were 1,000, 2,000 and 4,000 people and 100 QTLs was distributed on 10 chromosomes. Predictions were done by RKHS and BayesB models and mean comparisons were also made using the Tukey method.

Results: In all scenarios and with both methods, there was a significant difference between considering QTLs alone and with different percentages of noncausal SNPs in the model. However, in both methods, significant differences were observed between low percentages of noncausal markers and high percentages. In most scenarios, for example, when there were 10% noncausal markers in the model, the predictions were more accurate than when 20%, 40%, 60% and 100%. The accuracy of genomic predictions increased with the increasing of heritability and the number of individuals in the reference population. In the case of low heritability scenarios, the percentage of increment of predictive accuracy was more than higher heritability with increasing of reference population. The

Article history:

Received: 09/01/2025

Revised: 20/01/2025

Accepted: 03/02/2025

Keywords:

SNP

QTL

RKHS and BayesB models

mean of accuracy of evaluations using BayesB method compared RKHS, was significantly different in fitting QTLs in the model and different ratios of noncausal markers in most scenarios except for a few cases. In general, in both methods, the best predictions were made when using the highest heritability and population size and only entering QTLs in the models.

Conclusion: According to these findings, identifying causal sites and incorporating them into genomic assessments increases the rate of analysis and reduces the cost of genotyping. It can also enhance the genetic gain of important economic traits in genomic evaluations by enhancing predictive accuracy.

Cite this article: Saleh, M., Farhadi, A., Abdollahi-Arpanahi, R., Gholizadeh, M., Amiri-Roodbar, M. (2026). The effect of using different levels of noncausal markers along with causal loci and different genetic architectures on the accuracy of predicting genomic values in livestock. *Journal of Ruminant Research*, 14(1), 143-157.



© The Author(s)



10.22069/ejrr.2026.23949.2021

Publisher: Gorgan University of Agricultural Sciences and Natural Resources

تأثیر استفاده از سطوح مختلف نشانگرهای خنثی به همراه جایگاه‌های سببی و معماری‌های ژنتیکی مختلف بر صحت پیش‌بینی ارزش‌های اصلاحی ژنومی در دام‌های اهلی

معصومه صالح^۱، ایوب فرهادی^{۲*}، رستم عبداللهی آرپناهی^۳، محسن قلی‌زاده^۴، محمود امیری رودبار^۴

^۱ دانشجوی دکتری رشته ژنتیک و اصلاح نژاد دام، دانشگاه کشاورزی و منابع طبیعی ساری، دانشکده علوم دامی و شیلات، گروه علوم دامی،
^۲ (نویسنده مسئول) دانشیار رشته ژنتیک و اصلاح نژاد دام و مدیر گروه علوم دامی، دانشگاه کشاورزی و منابع طبیعی ساری، دانشکده علوم دامی و شیلات،
رایانامه: ayoub_farhadi@ymail.com
^۳ استادیار رشته ژنتیک و اصلاح نژاد دام، دانشگاه ابوریحان، گروه علوم دامی، محقق گروه علوم دامی دانشگاه فلوریدا،
^۴ استادیار پژوهشی رشته ژنتیک و اصلاح نژاد دام، مرکز تحقیق و توسعه کشاورزی و منابع طبیعی صفا آباد دزفول،

اطلاعات مقاله چکیده

نوع مقاله:	سابقه و هدف: مهم‌ترین سازه‌هایی که بر صحت پیش‌بینی ژنومی تأثیر می‌گذارند شامل تعداد نشانگرها، روش آماری، فراوانی آلل نادر، وراثت‌پذیری و معماری ژنتیکی می‌باشند که هر یک تحت شرایط ویژه‌ای اثراتشان متفاوت است. هدف از این پژوهش، بررسی اثر اضافه کردن چندشکلی‌های تک نوکلئوتیدی خنثی (غیرسببی) به مدل آماری بر صحت ارزیابی ارزش اصلاحی با در نظر گرفتن سازه‌های مختلف جمعیت و معماری‌های ژنتیکی مختلف بوده است.
تاریخ دریافت: ۱۰۱۴۰۴/۱۹	
تاریخ ویرایش: ۱۴۰۴/۱۰/۳۰	
تاریخ پذیرش: ۱۴۰۴/۱۱/۱۴	
واژه‌های کلیدی:	مواد و روش‌ها: در این پژوهش تعداد ۱۰ کروموزوم به طول ۱۰۰ ساتی مورگان شبیه‌سازی شده و روی هر کروموزوم ۱۰۰۰ نشانگر چندشکلی تک نوکلئوتیدی دو آللی در نظر گرفته شده که به صورت تصادفی در طول ژنوم پراکنده شدند. در این پژوهش، واریانس فنوتیپی یک و وراثت‌پذیری‌ها بر طبق صفات با وراثت‌پذیری کم، متوسط و بالا به ترتیب ۰/۱، ۰/۲ و ۰/۵ در نظر گرفته شدند. در این شبیه‌سازی، جمعیت‌های با اندازه ۱۰۰۰، ۲۰۰۰ و ۴۰۰۰ فرد مورداستفاده قرار گرفت و تعداد ۱۰۰ جایگاه صفت کمی روی ۱۰ کروموزوم پراکنده شد. پیش‌بینی‌ها با دو روش کرنل بازتولید فضای هیلبرت و بیز B و مقایسات میانگین با استفاده از روش توکی انجام شد.
چندشکلی تک نوکلئوتیدی جایگاه صفات کمی کرنل بازتولید فضای هیلبرت و بیز B	یافته‌ها: در همه‌ی سناریوها و با هر دو روش اختلاف معنی‌داری بین در نظر گرفتن جایگاه‌های صفات کمی به‌تنهایی و همراه با درصد‌های مختلف چندشکلی‌های تک نوکلئوتیدی غیرسببی در مدل مشاهده شد. البته در هر دو روش اختلافات معنی‌داری نیز بین درصد‌های پایین نشانگرهای خنثی و درصد‌های بالا مشاهده شد. در بیشتر سناریوها مثلاً زمانی که ۱۰ درصد نشانگرهای خنثی در مدل وجود دارند، پیش‌بینی‌ها صحت بالاتری دارند نسبت به زمانی که ۲۰ درصد، ۴۰ درصد، ۶۰ درصد و ۱۰۰ درصد از نشانگرهای خنثی وارد مدل می‌شوند. با افزایش وراثت‌پذیری و تعداد افراد جمعیت مرجع صحت پیش‌بینی‌های ژنومی افزایش یافت. در

وراثت‌پذیری‌های پایین، درصد افزایش صحت پیش‌بینی با افزایش جمعیت مرجع بیشتر بود. میانگین صحت ارزیابی با استفاده از روش بیز B به‌جز چند مورد، در زمان قرار دادن جایگاه-های صفات کمی در مدل و نسبت‌های مختلف نشانگرهای خنثی به‌طور معنی‌داری در بیشتر سناریوها از روش کرنل بازتولید فضای هیلبرت بیشتر بود. به‌طورکلی در هر دو روش، بهترین پیش‌بینی‌ها در زمان استفاده از بالاترین وراثت‌پذیری و اندازه جمعیت و فقط واردکردن جایگاه-های صفات کمی در مدل‌ها رخ داد.

نتیجه‌گیری: با توجه به این یافته‌ها، شناسایی جایگاه‌های سببی و واردکردن آن‌ها در ارزیابی‌های ژنومی باعث افزایش سرعت آنالیز و کاهش هزینه‌های تعیین ژنوتیپ می‌گردد و همچنین می‌تواند در ارزیابی‌های ژنومیکی بازده ژنتیکی صفات مهم اقتصادی را از طریق تقویت صحت پیش‌بینی افزایش دهد.

استناد: صالح، معصومه؛ فرهادی، ایوب؛ عبداللهی آرپناهی، رستم؛ قلی‌زاده، محسن؛ امیری رودبار، محمود. (۱۴۰۵). تأثیر استفاده از سطوح مختلف نشانگرهای خنثی به همراه جایگاه‌های سببی و معماری‌های ژنتیکی مختلف بر صحت پیش‌بینی ارزش‌های اصلاحی ژنومی در دام‌های اهلی. *پژوهش در نشخوارکنندگان*، ۱۴(۱)، ۱۵۷-۱۴۳.



10.22069/ejrr.2026.23949.2021

© نویسندگان

ناشر: دانشگاه علوم کشاورزی و منابع طبیعی گرگان



مقدمه

پیش‌بینی ژنومی یعنی پیش‌بینی و انتخاب بر پایه‌ی اطلاعات حاصل از نشانگرهای مولکولی با یا بدون استفاده از اطلاعات شجره، سودآوری را با افزایش پیشرفت ژنتیکی ناشی از کاهش فاصله نسلی، کاهش هزینه‌ها و افزایش صحت انتخاب، افزایش می‌دهد (Meuwissen و همکاران، ۲۰۰۱).

ارزیابی ژنومی صفات پیچیده در کشاورزی، دام و انسان با استفاده از مدل‌های مختلف خطی، غیرخطی و روش‌های یادگیری ماشین با استفاده از اطلاعات چندشکلی تک نوکلئوتیدی^۱ به‌عنوان متغیرهای پیش‌بینی کننده صورت می‌گیرد. از جایگاه‌های صفات کمی می‌توان در مدل برای به دست آوردن یک اثر ژنتیکی افزایشی که اثر تجمعی نشانگرهای ژنتیکی در فوتوتیپ‌های صفت را به دست می‌دهد، استفاده کرد. خصوصیات مختلف جایگاه‌های صفات کمی می‌تواند روی توانایی پیش‌بینی و قدرت تشخیص مجموعه‌ی نشانگری با استفاده از مدل‌های مختلف تأثیر بگذارد، از جمله نسبت واریانس کل ژنتیکی که توسط جایگاه‌های صفات کمی مورد نظر توجیه می‌شود، تعداد نواحی غیرسببی^۲ که در ناحیه‌ی جایگاه‌های صفات کمی وجود دارد و معماری ژنتیکی جایگاه‌های صفات کمی (مانند اینکه واریانت‌های سببی به صورت پراکنده یا دسته‌بندی شده^۳ روی ژنوم قرار دارند). علاوه بر این، مدل ممکن است تحت تأثیر سازه‌های صفت-ویژه یا جمعیت مانند وراثت‌پذیری صفت و تعداد مشاهدات قابل دسترس برای آزمون نیز قرار گیرد (Sarup و همکاران، ۲۰۱۶).

بر طبق مطالعه‌ی دیگری نیز مهم‌ترین سازه‌هایی که بر صحت پیش‌بینی تأثیر می‌گذارند، شامل تعداد

نشانگرها، روش آماری، فراوانی آلل نادر^۴، وراثت‌پذیری و معماری ژنتیکی می‌باشند که هر یک تحت شرایط ویژه‌ای اثراتشان متفاوت است. برای مثال، در مورد صفاتی که تحت تأثیر ژن‌های کمی قرار دارند افزایش تعداد نشانگرها باعث کاهش صحت می‌گردد و بالعکس. یا مثلاً روش بهترین پیش‌بینی ناریب خطی ژنومی^۵ در مقایسه با روش‌های بیز در مورد صفات با وراثت‌پذیری کم بهتر عمل کرده و صفات‌های بالاتری به دست می‌دهد و همچنین روش بهترین پیش‌بینی ناریب خطی ژنومی کمتر تحت تأثیر میزان فراوانی آلل نادر قرار می‌گیرد، در حالیکه روش بیزلاسو در فراوانی‌های آلل نادر بسیار کم نیز صفات‌های بسیار کمی می‌دهد (Zhang و همکاران، ۲۰۱۹).

روش‌های بیزین^۶ یک توزیع پیشین را برای توزیع اثرات جایگاه‌های صفات کمی و توزیع دیگری را برای تعداد جایگاه‌های صفات کمی در نظر می‌گیرند. توزیع اثرات جایگاه‌های صفات کمی برای بسیاری از صفات کمی نامشخص است (Meuwissen و همکاران، ۲۰۰۱) ولی در بسیاری از مطالعات نشان داده شده است که در نظر گرفتن توزیع گاما نسبت به توزیع نرمال صحت‌های بالاتری به دست می‌دهد (Goddard، ۲۰۰۹).

اینکه چگونه این روش‌ها و سازه‌ها بر انتخاب ژنومی تأثیر می‌گذارد، هنوز کاملاً مشخص نیست و یک عدم اطمینان را در انتخاب روش برای جمعیت یا صفت خاصی را موجب می‌شود. به‌هرحال فرض بر این است که کاربرد نسبی روش‌های ارزیابی ژنومی به ساختار ژنومی جمعیت و همچنین معماری ژنتیکی صفت بستگی دارد؛ بنابراین، هدف اصلی این مطالعه تأثیر سطوح متفاوت نشانگرهای غیرسببی بر صحت

⁴ Minor allele frequency (MAF)

⁵ Genomic best linear unbiased prediction (GBLUP)

⁶ Bayesian methods

¹ Single nucleotide polymorphism (SNP)

² Non causal

³ Cluster

در این مطالعه تعداد ۱۰ کروموزوم به عنوان نمونه‌ای از ژنوم گاو که هر یک ۱۰۰ سانتی‌مورگان طول داشتند، شبیه‌سازی شد. به دلیل کاهش زمان محاسباتی این تعداد کروموزوم در نظر گرفته شد. روی هر کروموزوم ۱۰۰۰ نشانگر چندشکلی تک نوکلئوتیدی دو آللی در نظر گرفته شده که به صورت تصادفی در طول ژنوم پراکنده شدند. در این مطالعه واریانس فنوتیپی یک در نظر گرفته شده و سه صفت با وراثت‌پذیری کم (۰/۱۰)، متوسط (۰/۲۰) و بالا (۰/۵۰) شبیه‌سازی شدند. نرخ جهش نشانگر در این شبیه‌سازی $2/5 \times 10^{-5}$ در نظر گرفته شده است.

شبیه‌سازی جایگاه‌های صفات کمی: خروجی فایل ژنوتیپ حاصل از QMSim به برنامه R نسخه‌ی 3.5.2 منتقل شد و از بین تمام نشانگرهای موجود ۱۰۰ چندشکلی تک نوکلئوتیدی با استفاده از توزیع نرمال انتخاب شد و اثراتی با میانگین صفر و انحراف معیار (وراثت‌پذیری بخش بر k) به آن‌ها اختصاص داده شد.

$$sd = \sqrt{\frac{h2_snp}{k}}$$

ضریب k حاصل جمع واریانس جایگاه‌های صفات کمی موردنظر است.

درنهایت تعداد ۱۰ جایگاه‌های صفات کمی دو آللی که به صورت تصادفی بر روی هر کروموزوم قرار گرفته بودند، شبیه‌سازی شد. فنوتیپ‌ها با جمع اثرات جایگاه‌های صفات کمی هر فرد به علاوه مقدار باقی‌مانده که از توزیع نرمال با میانگین صفر و واریانس یک شبیه‌سازی شد؛ محاسبه گردید. تمامی سناریوها ۱۰ بار تکرار شدند. از ۵ نسل آخر در جمعیت ثانویه چهار نسل (۱۰۱ تا ۱۰۴) به‌عنوان گروه مرجع و نسل آخر (۱۰۵) به‌عنوان گروه تائید انتخاب شدند.

ارزیابی‌های ژنومی با دو روش بیز B و کرنل بازتولید فضای هیلبرت^۱ تحت وراثت‌پذیری و اندازه جمعیت‌های مختلف بود. وارد نمودن و خارج نمودن اطلاعات بیولوژیکی مانند جایگاه‌های صفات کمی در مدل‌های پیش‌بینی با استفاده از مدل‌های ناپارامتری مانند روش‌های کرنلی، راهکار نوینی است که در این پژوهش موردبررسی قرار گرفت.

مواد و روش‌ها

شبیه‌سازی جمعیت و ژنوم: شبیه‌سازی جمعیت با استفاده از نرم‌افزار QMSim 1.10,2013 (Sargolzaei و Schenkel، ۲۰۰۹) انجام شده که در ابتدا تعداد ۱۰۰۰ حیوان به‌عنوان جمعیت پایه و به‌منظور ایجاد سطوح مناسبی از عدم تعادل لینکاژ^۲ و همچنین ایجاد تعادل جهش-رانس شبیه‌سازی شد که این ساختار جمعیتی از طریق یک گلوگاه^۳ ژنتیکی پس از ۸۰ نسل به ۱۰۰ رأس رسیده و به مدت ۱۸ نسل از طریق آمیزش تصادفی تداوم یافت و طی یک نسل به ۵۰۰ رأس افزایش یافته و درنهایت در نسل ۱۰۰ به ۵۵۰ رأس افزایش یافته است. در نسل آخر، از ۵۵۰ رأس، ۴۰ رأس به‌عنوان جنس نر و ۵۱۰ رأس از جنس ماده برای ادامه آمیزش‌های تصادفی در نظر گرفته شد. ۵ نسل دیگر آمیزش تصادفی به‌منظور ایجاد جمعیت ثانویه صورت گرفت که برای شبیه‌سازی جمعیت با اندازه ۱۰۰۰ فردی، ۱۰۰ ماده و ۱۰ نر با ۲ فرزند برای هر ماده، برای جمعیت مرجع ۲۰۰۰ فردی، ۲۰۰ ماده و ۲۰ نر با ۲ فرزند برای هر ماده و برای جمعیت ۴۰۰۰ فردی، ۴۰۰ ماده و ۴۰ رأس نر با ۲ فرزند برای هر ماده در هر نسل شبیه‌سازی شد. درنهایت میزان نامتعادلی لینکاژ به ۰/۱۵ تا ۰/۳ بین جفت نشانگرها رسید.

¹ Reproducing kernel Hilbert space (RKHS)

² Linkage disequilibrium (LD)

³ Bottleneck

آنالیز آماری و پیش‌بینی ارزش‌های اصلاحی:

آنالیزها ابتدا به‌تنهایی با استفاده از جایگاه‌های سببی یا

جایگاه‌های صفات کمی انجام گرفته و سپس با اضافه

کردن به ترتیب ۱۰، ۲۰، ۴۰، ۶۰ و ۱۰۰ درصد از

چندشکلی‌های تک نوکلئوتیدی خنثی به مدل

موردبررسی قرار گرفتند. دو حالت دیگر یکی با

حذف کامل جایگاه‌های صفات کمی از مدل و یکی

با حذف ۵۰ درصد جایگاه‌های صفات کمی در

شرایطی که همه چندشکلی‌های تک نوکلئوتیدی

خنثی در مدل بودند نیز انجام شده است. نهایتاً تأثیر

واردکردن اطلاعات اضافی به مدل آماری با استفاده از

برآورد وراثت‌پذیری و توانایی پیش‌بینی در بین

سناریوها با هم مقایسه شدند.

مدل آماری به‌صورت زیر بود (Sarup و همکاران،

۲۰۱۶)

$$y' = \mu + Zg + e$$

که g بردار مقادیر ژنومیکی مربوط به همه‌ی

نشانگرهای ژنتیکی است.

در زمان استفاده از هر دو جایگاه‌ها (سببی و خنثی)

مدل آماری به‌صورت زیر تعریف می‌شود.

$$\hat{y} = \mu + Zf + Zr + e$$

که \hat{y} بردار فنوتیپ‌های تصحیح‌شده، μ میانگین کل،

Z ماتریس طرحی که مشاهدات را به اثرات نشانگرها

ربط می‌دهد، f بردار مقادیر ژنومیکی توضیح داده‌شده

با جایگاه‌های ژنتیکی سببی است، r بردار مقادیر

ارزش‌های اصلاحی ژنومیکی مربوط به نشانگرهای

ژنتیکی گروه باقی‌مانده است و e بردار

باقی‌مانده‌هاست.

در روش بیز B بخشی از نشانگرها که اطلاعاتی

در مورد فنوتیپ فراهم می‌آورند در مدل قرار می‌-

گیرند. علیرغم معایب فرمولی، تقریباً قابل قبول‌ترین

مدل رایج می‌باشد. ترکیبی از توزیع پیشین نرمال با

میانگین صفر و واریانس σ_j^2 برای اثرات نشانگری

$g_j | \sigma_j^2 \sim N(0, \sigma_j^2)$ فرض می‌شود که توزیع صفر و

کای مربع وارون با پارامترهای (v, S^2) به نسبت π و

$1 - \pi$ برای واریانس‌ها در نظر گرفته می‌شود:

$$\begin{cases} \sigma_j^2 = 0 & \pi \\ \sigma_j^2 \sim \chi^{-2}(v, S) & 1 - \pi \end{cases}$$

این مدل با استفاده از نرم‌افزار R و نمونه‌گیری-

های گیبس با طول زنجیره نمونه‌گیری گیبس ۲۰۰۰۰

تکرار، تعداد تکرارهای قلق‌گیری ۵۰۰۰ و فاصله

نمونه‌گیری ۵ انجام شده است.

Gianola و همکاران (۲۰۰۶) یک روش نیمه

پارامتری به نام کرنل بازتولید فضای هیلبرت را در

ارزیابی ژنومی به‌عنوان جایگزینی برای رگرسیون

چندشکلی تک نوکلئوتیدی پیشنهاد کردند. انتظار این

بود که این روش‌ها برای گرفتن الگوهای تعامل

پیچیده که ممکن است پاسخ دادن به آن‌ها در

مدل‌های خطی دشوار است مناسب‌تر باشند

(Goddard, ۲۰۰۹). در پژوهش حاضر از این روش

استفاده گردید که برای اجرای این مدل از بسته

نرم‌افزاری BGLR استفاده می‌شود (Perez و de los

Campos, ۲۰۱۴).

در رویکرد بیزی مدل‌های آماری مشابه با رویکرد

REML هستند با این تفاوت که اثرات تصادفی

ژنتیکی افزایشی، از ضرب ماتریس خویشاوندی کرنل

خطی (K) در α حاصل می‌شوند. در مدل کرنل

بازتولید فضای هیلبرت، K یک ماتریس کرنل $n \times n$

بوده که به‌وسیله متغیرهای پیوسته چندشکلی تک

نوکلئوتیدی نمایه شده است و $K \sim XX^T$ بوده که در

آن X : ماتریس ژنوتیپ چندشکلی‌های تک

نوکلئوتیدی است. K از ماتریس X مرکزی شده و

استاندارد به دست آمد، سپس طبق پیشنهاد

VanRaden (۲۰۰۸) و Yang و همکاران (۲۰۱۰) بر

تعداد چندشکلی تک نوکلئوتیدی تقسیم‌شده است. α

مقایسه میانگین صحت‌های پیش‌بینی: در این مطالعه برای مقایسه میانگین صحت‌های ارزیابی از روش توکی استفاده شده است. برای انجام این روش از پکیج Agricola در نرم‌افزار R استفاده گردید.

نتایج و بحث

صحت پیش‌بینی ارزش اصلاحی برای صفات کمی در وراثت‌پذیری‌ها، اندازه‌های جمعیت و تراکم نشانگری مختلف در پلات‌های ۱، ۳، ۵ و ۷ آورده شده است. تقریباً در همه‌ی وراثت‌پذیری-جمعیت‌ها بین در نظر گرفتن جایگاه‌های صفات کمی به‌تنهایی (۰ درصد) و در نظر گرفتن جایگاه‌های صفات کمی همراه با درصد‌های مختلف چندشکلی‌های تک نوکلئوتیدی غیرسببی در مدل تفاوت معنی‌داری در صحت ارزیابی وجود داشت که نشان‌دهنده این است که فقط وجود چندشکلی‌های تک نوکلئوتیدی سببی برای پیش‌بینی فنوتیپ یا ارزش اصلاحی در جمعیت تأیید کافی است و بالاترین صحت ارزیابی را به دست می‌دهد. این می‌تواند به این دلیل باشد که هر چه بار اضافی به مدل تحمیل می‌شود، مدل سخت‌تر برآزش شده و پیچیدگی مدل افزایش می‌یابد و بنابراین انتظار بر این است که این بار اضافی باعث دور شدن پیش‌بینی‌ها از مقدار واقعی و بنابراین کاهش صحت پیش‌بینی گردد. مطالعات زیادی این نتیجه را تأیید کرده‌اند که اضافه نمودن متغیرهای اضافی که اطلاعات یا اثر خاصی ندارند به مدل باعث کاهش صحت ارزیابی می‌گردد (Amiri Roudbar و همکاران، ۲۰۲۰؛ Sarup و همکاران، ۲۰۱۶).

با استناد به مطالعات قبلی تعداد جایگاه‌های صفات کمی در همه سناریوها یکسان در نظر گرفته شده و تنها توزیع اثرات بین آن‌ها متفاوت بود. البته در اضافه کردن نشانگرهای خنثی به مدل یک‌بار جایگاه‌های صفات کمی به‌طورکلی از مدل حذف شد

بردار ضرایب رگرسیون کرنل بازتولید فضای هیلبرت است که تابع درست‌نمایی را به‌صورت زیر حداکثر می‌کند:

$$l(\alpha|\lambda) = (y-K\alpha)'(y-K\alpha) + \lambda\alpha'K\alpha$$

در رابطه بالا $(\alpha \sim N(0, K^{-1}\sigma_g^2))$ ، واریانس ژنومی افزایشی توجیه شده به وسیله نشانگرها و λ : یک فاکتور تنظیم‌کننده است. مقادیر ارزش‌های ژنتیکی افزایشی از رابطه $g = K\alpha$ به دست می‌آید. در روش کرنل بازتولید فضای هیلبرت ابعاد معادلات از $m \times m$ یعنی صدها هزار چندشکلی تک نوکلئوتیدی به تعداد $n \times n$ یعنی برابر با تعداد مشاهدات کاهش می‌یابد. همان‌طور که گفته شد از روش بی‌زی برای برآورد اجزای واریانس ژنومی با استفاده از تکنیک نمونه‌گیری گیبس و مدل آماری کرنل بازتولید فضای هیلبرت در محیط R استفاده شد. در پژوهش حاضر، طول زنجیره نمونه‌گیری گیبس ۲۰۰۰۰ تکرار، تعداد تکرارهای قلق‌گیری ۵۰۰۰ و فاصله نمونه‌گیری ۵ در نظر گرفته شد.

معیارهای ارزیابی

برای بررسی توانایی پیش‌بینی سناریوها از دو معیار همبستگی و MSE (میانگین مربعات خطا) بین ارزش اصلاحی واقعی و پیش‌بینی‌شده استفاده شد که هرکدام به‌صورت زیر محاسبه گردید:

$$r_{a,\bar{y}} = \frac{\sigma_{a,\bar{y}}}{\sigma_a \sigma_{\bar{y}}}$$

$\sigma_{a,\bar{y}}$ کوواریانس بین اثرات افزایشی و فنوتیپ پیش‌بینی‌شده است و σ_a انحراف معیار اثرات افزایشی و $\sigma_{\bar{y}}$ انحراف معیار فنوتیپ پیش‌بینی‌شده است که در اینجا معادل ارزش اصلاحی پیش‌بینی‌شده می‌باشد.

$$MSE = \frac{(a - \bar{y})^2}{n}$$

a اثرات افزایشی واقعی و \bar{y} فنوتیپ پیش‌بینی‌شده می‌باشد.

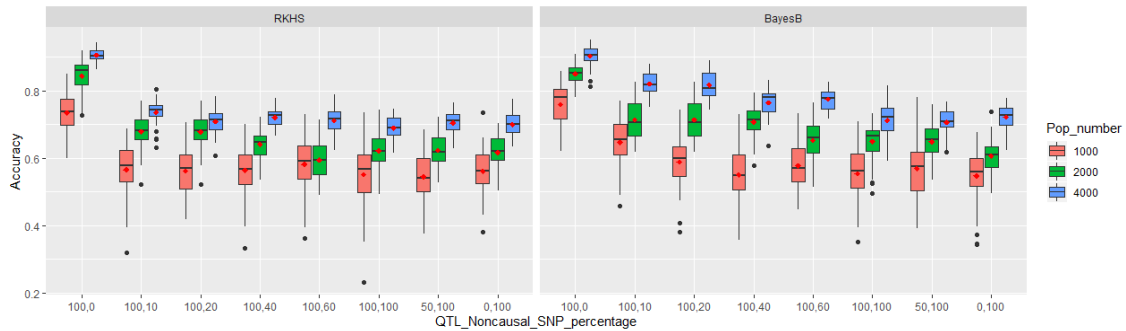
ارزیابی در وراثت‌پذیری ۰/۱ و در سه جمعیت موردبررسی در نسبت‌های مختلف تراکم نشانگری در پلات مربوط به شکل ۱ نشان داده شده است. میانگین صحت با افزایش جمعیت افزایش یافته و با افزایش تراکم نشانگری کاهش می‌یابد. هم‌چنین همان‌طور که در شکل ۲ مشاهده می‌شود، با افزایش اندازه جمعیت میانگین مربعات خطای پیش‌بینی کاهش یافته و با افزایش تراکم نشانگری افزایش می‌یابد.

بنابر گزارشات پیشین علاوه بر وراثت‌پذیری، عدم تعادل بین نشانگرها، رابطه خویشاوندی بین جمعیت مرجع و تأیید و هم‌چنین تعداد افراد موردمطالعه بر صحت پیش‌بینی‌ها تأثیرگذار هستند (Foroutanifar و همکاران، ۲۰۱۲). در مورد صفاتی با وراثت‌پذیری کم، افزایش اندازه جمعیت موجب افزایش صحت ارزیابی می‌شود. در واقع اختلاف معنی‌داری بین جمعیت ۱۰۰۰ و ۴۰۰۰ و هم‌چنین ۱۰۰۰ و ۲۰۰۰ وجود دارد ($P < 0.001$). درصد افزایش صحت در مورد صفات با وراثت‌پذیری کم با افزایش جمعیت بسیار بیشتر از صفات با وراثت‌پذیری متوسط و بالاست. در این وراثت‌پذیری بیشترین اختلافات معنی‌دار بین دو روش بیز B و کرنل بازتولید فضای هیلبرت در مورد جمعیت ۴۰۰۰ است که در همه نسبت‌ها به جز حالتی که درصد نشانگرهای خنثی در مدل صفر است، اختلاف معنی‌داری بین دو روش مشاهده می‌شود. البته در جمعیت‌های ۱۰۰۰ و ۲۰۰۰ نیز در برخی نسبت‌ها اختلافات معنی‌داری بین دو روش وجود دارد.

و یک‌بار نیمی از آن‌ها همراه با نشانگرهای خنثی در مدل آورده شد که در هر دو حالت نسبت به حالتی که همه جایگاه‌های صفات کمی در مدل بود، صحت ارزیابی کاهش یافته بود. برابر مطالعات انجام شده، افزایش تعداد جایگاه‌های کنترل‌کننده صفات کمی و هم‌چنین تفاوت توزیع آثار آن‌ها تأثیری بر صحت ارزیابی‌های ژنومیکی ندارد (Atashi و Gorgani-Firoozjah، ۲۰۱۵؛ Mahmoudi و همکاران، ۲۰۰۵؛ Foroutanifar، ۲۰۱۶). البته نظراتی مخالف این هم وجود دارد. برخلاف مطالعات قبلی نشان داده شد که با افزایش تعداد جایگاه‌های صفات کمی صحت افزایش یافته و میزان این افزایش بستگی به وراثت‌پذیری صفت دارد (Naderi و همکاران، ۲۰۱۲).

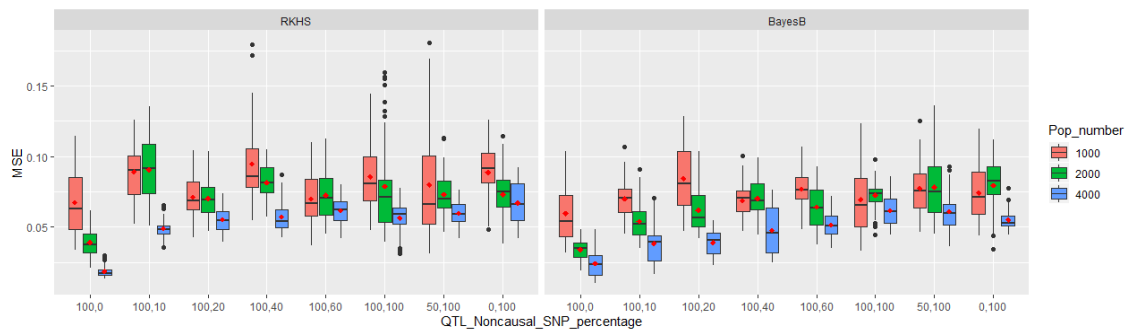
همان‌طور که بیان شد جایگاه‌های صفات کمی به دو صورت بزرگ اثر و کوچک اثر در نظر گرفته شدند؛ یعنی دو درصد جایگاه‌های صفات کمی ۳۰ درصد واریانس را کنترل کردند که طبق مطالعات پیشین اثر مثبتی را برافزایش صحت ایفا کرده است. مطالعه Coster و همکاران (۲۰۱۰) نشان داد که زمانی که ۱۰ درصد جایگاه‌های صفات کمی، ۹۰ درصد واریانس صفت را کنترل می‌کردند با روش بیز B نسبت به حالت واریانس یکسان برای جایگاه‌های صفات کمی صحت بالاتری برای پیش‌بینی‌ها به دست آمد. البته تعداد محدود جایگاه‌های صفات کمی در مطالعات شبیه‌سازی، تأثیر مثبتی بر صحت پیش‌بینی‌ها داشته است (Calus و همکاران، ۲۰۰۸؛ Solberg و همکاران، ۲۰۰۸؛ Meuwissen، ۲۰۰۹).

صفات با وراثت‌پذیری کم: بسیاری از صفات کمی در دام‌های اهلی از جمله صفات تولیدمثلی، وراثت‌پذیری پایینی دارند. در این مطالعه صحت



شکل ۱: صحت پیش‌بینی در مقابل تراکم نشانگری (اضافه نمودن جایگاه‌های غیرسببی) در وراثت‌پذیری ۰/۱ با هر دو مدل بیز و کرنل بازتولید فضای هیلبرت. اعداد روی محور X اولین عدد درصد وجود جایگاه‌های صفات کمی در مدل و دومین عدد نشان‌دهنده درصد نشانگرهای غیرسببی است که هر بار به مدل اضافه شده است. 1000 اندازه جمعیت ۱۰۰۰ فرد، 2000 اندازه جمعیت ۲۰۰۰ فرد و 4000 اندازه جمعیت ۴۰۰۰ فرد را نشان می‌دهد. نقاط قرمز روی هر جعبه مربوط به میانگین گروه موردنظر می‌باشد.

Figure 1. The accuracy of prediction against marker density (addition of non-causal SNPs) for heritability 0.1 with both Base B and RKHS models. The numbers on the x-axis are the first percentage of the QTLs in the model and the second is the percentage of the non-causal SNPs added to the model each time. 1000 shows the population size of 1000 people, 2000 population size shows 2000 people and 4000 population size show 4000 people. The red dots on each box correspond to the average of the desired group.



شکل ۲: میانگین مربعات خطای پیش‌بینی در مقابل تراکم نشانگری (اضافه نمودن نشانگرهای غیرسببی) در وراثت‌پذیری ۰/۱ با هر دو مدل بیز و کرنل بازتولید فضای هیلبرت. اعداد روی محور X اولین عدد درصد وجود جایگاه‌های صفات کمی در مدل و دومین عدد نشان‌دهنده درصد نشانگرهای غیرسببی است که هر بار به مدل اضافه شده است. 1000 اندازه جمعیت ۱۰۰۰ فرد، 2000 اندازه جمعیت ۲۰۰۰ فرد و 4000 اندازه جمعیت ۴۰۰۰ فرد را نشان می‌دهد. نقاط قرمز روی هر جعبه مربوط به میانگین گروه موردنظر می‌باشد.

Figure 2. The mean square error of prediction against marker density (addition of non-causal SNPs) for heritability 0.1 with both Base B and RKHS models. The numbers on the x-axis are the first percentage of the QTLs in the model and the second is the percentage of the non-causal SNPs added to the model each time. 1000 shows the population size of 1000 people, 2000 population size shows 2000 people and 4000 population size show 4000 people. The red dots on each box correspond to the average of the desired group.

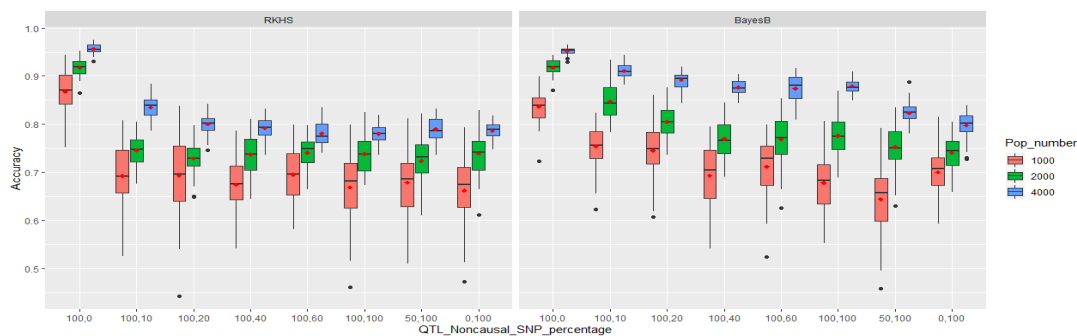
افزایش می‌یابد. اختلاف بین جمعیت‌های ۱۰۰۰ و ۴۰۰۰ به شدت معنی‌دار بود ($P < 0.001$) در حالیکه بین ۱۰۰۰ و ۲۰۰۰ یا ۲۰۰۰ و ۴۰۰۰ معنی‌دار نبوده است.

همانطور که در شکل ۴ مشاهده می‌شود با افزایش اندازه جمعیت میانگین مربعات خطای پیش‌بینی نیز کاهش یافته و با افزایش تراکم نشانگری افزایش می‌

صفات با وراثت‌پذیری متوسط: صفات تولید شیر، درصد چربی، ضریب تبدیل غذایی، وزن از شیرگیری و غیره وراثت‌پذیری متوسطی دارند. در پژوهش حاضر صحت‌های ارزیابی در وراثت‌پذیری ۰/۲ برای جمعیت‌ها و تراکم‌های مختلف نشانگری در شکل ۳ نشان داده شده است همان‌گونه که مشاهده می‌گردد در این حالت نیز با افزایش اندازه جمعیت صحت

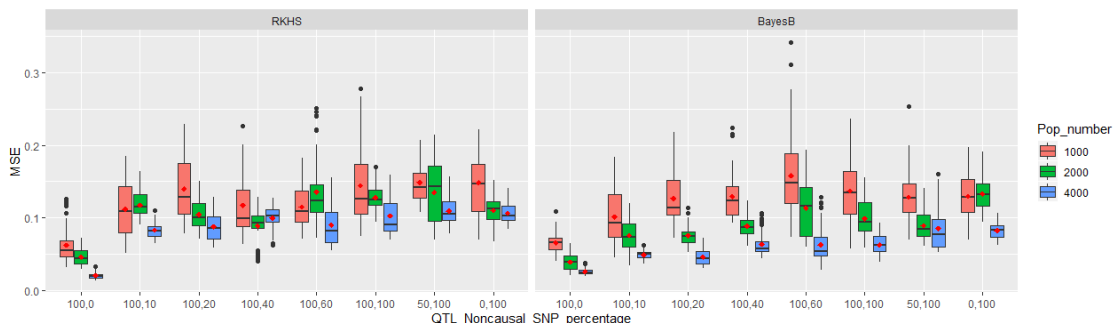
البته در جمعیت‌های ۱۰۰۰ و ۲۰۰۰ نیز در برخی نسبت‌ها اختلافات معنی‌داری بین دو روش وجود دارد.

یابد. در این وراثت‌پذیری بیشترین اختلافات معنی‌دار بین دو روش بیز B و کرنل بازتولید فضای هیلبرت در مورد جمعیت ۴۰۰۰ است که در همه نسبت‌ها اختلاف معنی‌داری بین دو روش مشاهده می‌شود.



شکل ۳: صحت پیش‌بینی در مقابل تراکم نشانگری (اضافه نمودن جایگاه‌های غیرسببی) در وراثت‌پذیری ۰/۲ در هر دو مدل بیز B و کرنل بازتولید فضای هیلبرت. اعداد روی محور X اولین عدد درصد وجود جایگاه‌های صفات کمی در مدل و دومین عدد نشان‌دهنده درصد نشانگرهای غیرسببی است که هر بار به مدل اضافه شده است. ۱۰۰۰ اندازه جمعیت ۱۰۰۰ فرد، ۲۰۰۰ اندازه جمعیت ۲۰۰۰ فرد و ۴۰۰۰ اندازه جمعیت ۴۰۰۰ فرد را نشان می‌دهد. خط تیره روی هر جعبه مربوط به میانگین گروه موردنظر می‌باشد.

Figure 3. The accuracy of prediction against marker density (addition of non-causal SNPs) for heritability 0.2 with both Base B and RKHS models. The numbers on the x-axis are the first percentage of the QTLs in the model and the second is the percentage of the non-causal SNPs added to the model each time. 1000 shows the population size of 1000 people, 2000 population size shows 2000 people and 4000 population size show 4000 people. The red dots on each box correspond to the average of the desired group.



شکل ۴: میانگین مربعات خطای پیش‌بینی در مقابل تراکم نشانگری (اضافه نمودن نشانگرهای غیرسببی) در وراثت‌پذیری ۰/۲ در هر دو مدل بیز B و کرنل بازتولید فضای هیلبرت. اعداد روی محور X اولین عدد درصد وجود جایگاه‌های صفات کمی در مدل و دومین عدد نشان‌دهنده درصد نشانگرهای غیرسببی است که هر بار به مدل اضافه شده است. ۱۰۰۰ اندازه جمعیت ۱۰۰۰ فرد، ۲۰۰۰ اندازه جمعیت ۲۰۰۰ فرد و ۴۰۰۰ اندازه جمعیت ۴۰۰۰ فرد را نشان می‌دهد. خط تیره روی هر جعبه مربوط به میانگین گروه موردنظر می‌باشد.

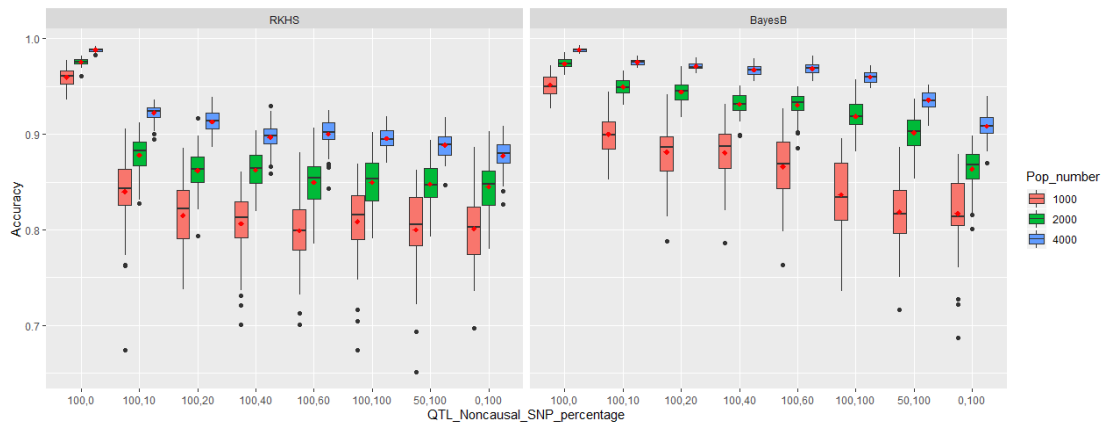
Figure 4. The mean square error of prediction against marker density (addition of non-causal SNPs) for heritability 0.2 with both Base B and RKHS models. The numbers on the x-axis are the first percentage of the QTLs in the model and the second is the percentage of the non-causal SNPs added to the model each time. 1000 shows the population size of 1000 people, 2000 population size shows 2000 people and 4000 population size show 4000 people. The red dots on each box correspond to the average of the desired group.

در وراثت‌پذیری ۰/۵ برای جمعیت‌ها و تراکم‌های مختلف نشانگری در شکل ۵ نشان داده شده است. در این حالت نیز مشاهده می‌گردد که با افزایش اندازه

صفات با وراثت‌پذیری بالا: صفاتی از جمله وزن تولد، وزن بلوغ، وزن تخم‌مرغ، ضریب تبدیل رشد و غیره وراثت‌پذیری بالایی دارند. صحت‌های ارزیابی

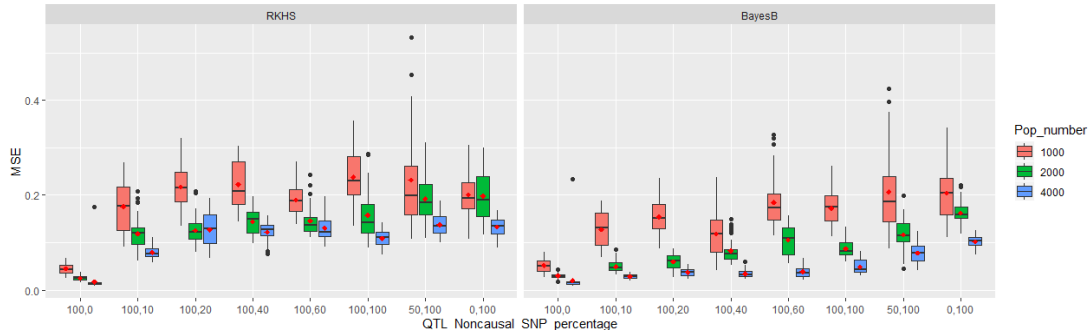
کمتر است. میانگین مربعات خطای پیش‌بینی نیز برای این وراثت‌پذیری در شکل ۶ نشان داده شده است.

جمعیت صحت افزایش می‌یابد. ولی این روند در مورد این صفات، کندتر از صفات با وراثت‌پذیری



شکل ۵: صحت پیش‌بینی در مقابل تراکم نشانگری (اضافه نمودن جایگاه‌های غیرسببی) در وراثت‌پذیری ۰/۵ در هر دو مدل بی‌ز و کرنل بازتولید فضای هیلبرت. اعداد روی محور X اولین عدد درصد وجود جایگاه‌های صفات کمی در مدل و دومین عدد نشان‌دهنده‌ی درصد نشانگرهای غیرسببی است که هر بار به مدل اضافه شده است. ۱۰۰۰ اندازه جمعیت ۱۰۰۰ فرد، ۲۰۰۰ اندازه جمعیت ۲۰۰۰ فرد و ۴۰۰۰ اندازه جمعیت ۴۰۰۰ فرد را نشان می‌دهد. نقاط قرمز روی هر جعبه مربوط به میانگین گروه موردنظر می‌باشد.

Figure 5: The accuracy of prediction against marker density (addition of non-causal SNPs) for heritability 0.5 with both Base B and RKHS models. The numbers on the x-axis are the first percentage of the QTLs in the model and the second is the percentage of the non-causal SNPs added to the model each time. 1000 shows the population size of 1000 people, 2000 population size shows 2000 people and 4000 population size show 4000 people. The red dots on each box correspond to the average of the desired group.



شکل ۶: میانگین مربعات خطای پیش‌بینی در مقابل تراکم نشانگری (اضافه نمودن نشانگرهای غیرسببی) در وراثت‌پذیری ۰/۵ در هر دو مدل بی‌ز و کرنل بازتولید فضای هیلبرت. اعداد روی محور X اولین عدد درصد وجود جایگاه‌های صفات کمی در مدل و دومین عدد نشان‌دهنده‌ی درصد نشانگرهای غیرسببی است که هر بار به مدل اضافه شده است. ۱۰۰۰ اندازه جمعیت ۱۰۰۰ فرد، ۲۰۰۰ اندازه جمعیت ۲۰۰۰ فرد و ۴۰۰۰ اندازه جمعیت ۴۰۰۰ فرد را نشان می‌دهد. نقاط قرمز روی هر جعبه مربوط به میانگین گروه موردنظر می‌باشد.

Figure 6: The mean square error of prediction against marker density (addition of non-causal SNPs) for heritability 0.5 with both Base B and RKHS models. The numbers on the x-axis are the first percentage of the QTLs in the model and the second is the percentage of the non-causal SNPs added to the model each time. 1000 shows the population size of 1000 people, 2000 population size shows 2000 people and 4000 population size show 4000 people. The red dots on each box correspond to the average of the desired group.

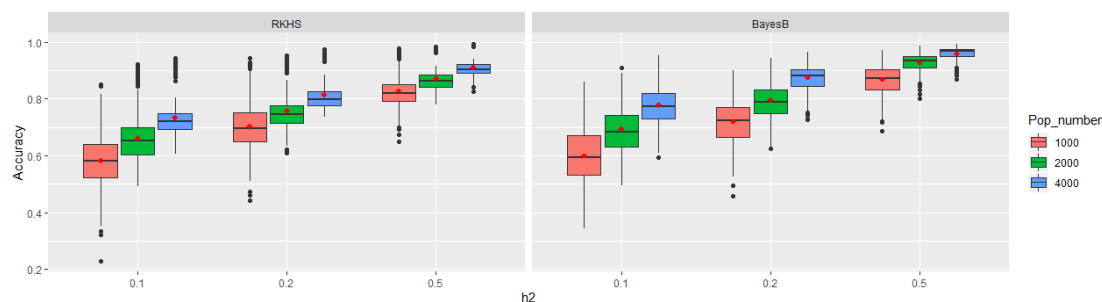
صحت ارزیابی‌ها افزایش یافته است. برخلاف نتایج این مطالعه، گزارشات زیادی مبنی بر افزایش صحت

Atefi و همکاران (۲۰۱۹) گزارش کردند که برای سه وراثت‌پذیری ۰/۱، ۰/۳ و ۰/۵ با افزایش جمعیت

صحت از ۰,۶۴ به ۰,۷ افزایش یافته و همچنین صحت پیش‌بینی به صورت غیرخطی تحت تأثیر تعداد جایگاه‌های صفات کمی هم قرار گرفت. در حقیقت، با افزایش وراثت‌پذیری صفت، صحت پیش‌بینی ارزش‌های اصلاحی ژنومیک بالا می‌رود و روند کاهش صحت پس از چند نسل از برآورد در جمعیت مرجع در وراثت‌پذیری‌های کم بیشتر است (Mohammadi Chamachar و همکاران، ۲۰۱۵). استفاده از روش کرنل بازتولید فضای هیلبرت به عنوان مدل پیش‌بینی در مطالعات زیادی برای صفات پیچیده قابلیت فوق‌العاده‌ای داشته است هرچند که در مورد سناریوهای تعریف‌شده در این مطالعه روش بیز B نتایج بهتری را داده است.

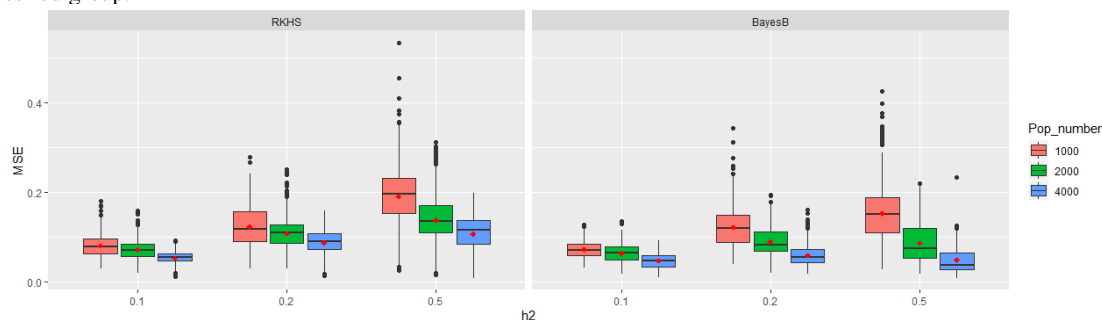
پیش‌بینی‌ها با افزایش تراکم نشانگری است (Wang و همکاران، ۲۰۱۷؛ Zhang و همکاران، ۲۰۱۷؛ Norman و همکاران، ۲۰۱۸؛ Hao و همکاران، ۲۰۱۹).

افزایش وراثت‌پذیری: در کل با افزایش وراثت‌پذیری در این مطالعه افزایش معنی‌داری در صحت پیش‌بینی رخ داده است. مطالعات زیادی وجود دارد که افزایش صحت در نتیجه افزایش وراثت‌پذیری را گزارش نموده‌اند (Calus، ۲۰۰۸؛ Solberg و همکاران، ۲۰۰۸؛ Luan و همکاران، ۲۰۰۹؛ Meuwissen، ۲۰۰۹؛ VanRaden و همکاران، ۲۰۱۱؛ Foroutanifar و همکاران، ۲۰۱۲؛ Xu و همکاران، ۲۰۱۹). بر طبق مطالعات Atefi و همکاران (۲۰۱۶)، افزایش وراثت‌پذیری تأثیر افزایشی مثبتی بر صحت پیش‌بینی -ها داشته است و با افزایش اندازه جمعیت مرجع



شکل ۷: صحت پیش‌بینی در مقابل وراثت‌پذیری. در هر دو مدل بیز B و کرنل بازتولید فضای هیلبرت.. اعداد روی محور x نشان‌دهنده‌ی وراثت‌پذیری‌های مختلف است. 1000 اندازه جمعیت ۱۰۰۰ فرد، 2000 اندازه جمعیت ۲۰۰۰ فرد و 4000 اندازه جمعیت ۴۰۰۰ فرد را نشان می‌دهد. نقاط قرمز روی هر جعبه مربوط به میانگین گروه موردنظر می‌باشد.

Figure 7: The accuracy of predictions against heritability in both Base B and RKHS models. The numbers on the x-axis represent different heritability. 1000 shows the population size of 1000 people, 2000 population size shows 2000 people and 4000 population size show 4000 people. The red dots on each box correspond to the average of the desired group.



شکل ۸: میانگین مربعات خطای پیش‌بینی در مقابل وراثت‌پذیری. در هر دو مدل بیز B و کرنل بازتولید فضای هیلبرت.. اعداد روی محور x نشان‌دهنده‌ی وراثت‌پذیری‌های مختلف است. 1000 اندازه جمعیت ۱۰۰۰ فرد، 2000 اندازه جمعیت ۲۰۰۰ فرد و 4000 اندازه جمعیت ۴۰۰۰ فرد را نشان می‌دهد. نقاط قرمز روی هر جعبه مربوط به میانگین گروه موردنظر می‌باشد.

Figure 8: The mean square error of prediction against heritability in both Base B and RKHS models. The numbers on the x-axis represent different heritability. 1000 shows the population size of 1000 people, 2000 population size shows 2000 people and 4000 population size show 4000 people. The red dots on each box correspond to the average of the desired group.

صحیح باعث کاهش صحت پیش‌بینی‌ها به‌طور معنی‌داری شد؛ بنابراین توصیه می‌شود با واردکردن اطلاعات بیولوژی مانند واریانت‌های سببی در مدل آماری و حذف چندشکلی‌های تک نوکلئوتیدی خنثی از ماتریس ژنوتایپی صحت پیش‌بینی بهبود داده شود ضمن این‌که در وراثت‌پذیری‌های مختلف روش بیز B نتایج بهتری از روش کرنل بازتولید فضای هیلبرت در پیش‌بینی‌ها دارد.

نتیجه‌گیری کلی

با توجه به مطالعات انجام‌شده معماری ژنتیکی صفات و پارامترهای جمعیت از جمله وراثت‌پذیری و اندازه جمعیت تأثیر بسزایی در صحت پیش‌بینی صفات کمی دارد. در این مطالعه افزایش اندازه جمعیت در گرو وراثت‌پذیری صفت بوده و در وراثت‌پذیری‌های پایین صحت پیش‌بینی با شدت بیشتری افزایش یافته است. اضافه نمودن واریانت‌های غیرسببی به مدل آماری

منابع

- Amiri Roudbar, M., Mohammadabadi, M.R., Ayatollahi Mehrgardi, A., Abdollahi-Arpanahi, R., Momen, M., Morota, G., Brito Lopes, F., Gianola, D. & Rosa, G.J.M. (2020). Integration of single nucleotide variants & whole-genome DNA methylation profiles for classification of rheumatoid arthritis cases from controls. *Heredity*, 124(5): 658-674.
- Atashi, H. & Gorgani-Firoozjah, N. (2015). Study of the effect of increasing the number of markers on the accuracy of genomic evaluation using rrBLUP method. *Iranian Journal of Animal Science*, 46: 195-200. (In Persian).
- Atefi, A., Shadparvar, A.A. & Ghavi Hossein-Zadeh, N. (2018). Accuracy of genomic prediction under different genetic architectures and estimation methods. *Iranian Journal of Applied Animal Science*, 8: 43-52.
- Atefi, A., Shadparvar, A.A. & Hossein-Zadeh, N.G. (2016). Comparison of whole genome prediction accuracy across generations using parametric and semi parametric methods. *Acta Scientiarum. Animal Sciences*, 38: 447-453.
- Calus, M.P.L., Meuwissen, T.H.E., de Roos, A.P.W. & Veerkamp, R.F. (2008). Accuracy of Genomic Selection Using Different Methods to Define Haplotypes. *Genetics*, 178: 553.
- Coster, A., Bastiaansen, J.W.M., Calus, M.P.L., van Arendonk, J.A.M. & Bovenhuis, H. (2010). Sensitivity of methods for estimating breeding values using genetic markers to the number of QTL and distribution of QTL variance. *Genetics Selection Evolution*, 42:9.
- Foroutanifar, S. (2016). Sensitivity of the accuracy of single-adjective and multi-adjective genomic predictions to the genetic architecture of traits. *Modern Genetics*, 11:391-398. (In Persian).
- Foroutanifar, S., Mehrabani-Yeganeh, H. & Moradi-Shahrbabak, M. (2012). Effects of heritability, number of individuals in training population and map density on accuracy of single and two traits estimated genomic breeding values. *Proceedings of 12th Iranian genetics congress*, 10: 21-23. (In Persian).
- Gianola, D., Fernando, R.L. & Stella, A. (2006). Genomic-assisted prediction of genetic value with semiparametric procedures. *Genetics*, 173:1761-1776.
- Goddard, M. (2009). Genomic selection: prediction of accuracy and maximisation of long-term response. *Genetica*, 136:245-57.
- Hao, Y., Wang, H., Yang, X., Zhang, H., He, C., Li, D., Li, H., Wang, G., Wang, J. & Fu, J. (2019). Genomic Prediction using Existing Historical Data Contributing to Selection in Biparental Populations: A Study of Kernel Oil in Maize. *The Plant Genome*, 12:1702-1712.

- Luan, T., Woolliams, J., Lien, S., Kent, M., Svendsen, M. & Meuwissen, T. (2009). The accuracy of genomic selection in Norwegian Red Cattle assessed by cross-validation. *Genetics*, 183: 1119-26.
- Mahmoudi, N.A.M., Ayatollahi mehrjerdi, A., Honarvar, M. & Ismaili Zadeh, A. (2015). Study of statistical distribution of QTL effects on the accuracy of genomic values estimated by Bayesian method. *Iranian Journal of Animal Science Research*, 7:356-363. (In Persian).
- Meuwissen, T.H.E. (2009). Accuracy of breeding values of 'unrelated' individuals predicted by dense SNP genotyping. *Genetics Selection Evolution*, 41:35.
- Meuwissen, T.H., Hayes, B.J. & Goddard, M.E. (2001). Prediction of total genetic value using genome-wide dense marker maps. *Genetics*, 157:1819-29.
- Mohammadi Chamachar, N. Hafezian, S.H., Honarvar, M. & Farhadi, A. (2015). The effect of heritability and the number of quantitative traits loci on the accuracy of estimating genomic value. *Journal of Ruminant Research*, 3:111-124. (In Persian).
- Naderi, R., Noshahri, A. & Amin Afshar, M. (2012). Effects of QTL variance on genomic selection accuracy. *Animal Science and Research Journal*, 81-91. (In Persian).
- Norman, A., Taylor, J., Edwards, J., & Kuchel, H. (2018). Optimizing Genomic Selection in Wheat: Effect of Marker Density, Population Size and Population Structure on Prediction Accuracy. *G3: Genes, Genomes, Genetics*, 8:2889-2899.
- Perez, P. & de los Campos, G. (2014). Genome-wide regression and prediction with the BGLR statistical package. *Genetics*, 198:483-95.
- Sargolzaei, M. & Schenkel, F.S. (2009). QMSim: a large-scale genome simulator for livestock. *Bioinformatics*, 25:680-681.
- Sarup, P., Jensen, J., Ostersen, T., Henryon, M & Sorensen, P. (2016). Increased prediction accuracy using a genomic feature model including prior information on quantitative trait locus regions in purebred Danish Duroc pigs. *BMC Genetics*, 17:11.
- Solberg, T.R., Sonesson, A.K., Woolliams, J.A. & Meuwissen, T.H. (2008). Genomic selection using different marker types and densities. *Journal of Animal Science*, 86:2447-54.
- VanRaden, P. M. (2008). Efficient methods to compute genomic predictions. *Journal of Dairy Science*, 91: 4414-4423.
- VanRaden, P.M., O'Connell, J.R., Wiggans, G.R. & Weigel, K.A. (2011). Genomic evaluations with many more genotypes. *Genetics Selection Evolution*, 43:10.
- Wang, Q., Yu, Y., Yuan, J., Zhang, X., Huang, H., Li, F. & Xiang, J. (2017). Effects of marker density and population structure on the genomic prediction accuracy for growth trait in Pacific white shrimp *Litopenaeus vannamei*. *BMC Genetics*, 18:45.
- Xu, L., Wang, Z., Zhu, B., Liu, Y., Li, H., Bordbar, F., Chen, Y., Zhang, L., Gao, X., Gao, H., Zhang, S., Xu, L. & Li, J. (2019). Theoretical Evaluation of Multi-Breed Genomic Prediction in Chinese Indigenous Cattle. *Animals*, 9:789.
- Yang, J., Benyamin, B., McEvoy, B. P., Gordon, S., Henders, A. K., Nyholt, D. R., & et al. (2010). Common SNPs explain a large proportion of the heritability for human height. *Genetics*, 42:565-569.
- Zhang, A., Wang, H., Beyene, Y., Semagn, K., Liu, Y., Cao, S., Cui, Z., Ruan, Y., Burgueño, J., San Vicente, F., Olsen, M., Prasanna, B.M., Crossa, J., Yu, H. & Zhang, X. (2017). Effect of Trait Heritability, Training Population Size and Marker Density on Genomic Prediction Accuracy Estimation in 22 bi-parental Tropical Maize Populations. *Frontiers in Plant Science*, 8:1916.
- Zhang, H., Yin, L., Wang, M., Yuan, X. & Liu, X. (2019). Factors Affecting the Accuracy of Genomic Selection for Agricultural Economic Traits in Maize, Cattle, and Pig Populations. *Frontiers in Genetics*, 10:189.