دانشگاه علوم کشاورزی و منابع طبیعی ساری
پژوهشهای تولیدات دامی

صحت انتخاب ژنومی روش‌های پارامتری و ناپارامتری با معماری‌های ژنتیکی افزایشی و غالبیت

یحیی محمدی[۱] و مرتضی ستایی مختاری[۲]

۱- استادیار، گروه علوم دامی، دانشگاه ایلام، (نویسنده مسوول: mohamadi_yahya@yahoo.com)
۲- استادیار، گروه علوم دامی، دانشگاه جیرفت

تاریخ دریافت: ۹۵/۱۱/۲۵ تاریخ پذیرش ۹۶/۳/۳۱

چکیده

در بیشتر مطالعات پیش‌بینی ژنومی برای برآورد ارزش اصلاحی ژنومی فقط از اثرات افزایشی در مدل استفاده می‌کنند. از طرفی در بیشتر صفات پیچیده یک منبع مهم تنوع، اثرات غالبیت بوده که لحاظ نمودن این اثر در مدل ممکن است باعث افزایش صحت پیش‌بینی‌های ژنومی گردند. در این مطالعه که با استفاده از شبیه‌سازی داده انجام شد، تاثیر مقادیر مختلف وراثت‌پذیری (۰/۱، ۰/۳ و ۰/۵) و نیز مقادیر مختلف نسبت واریانس غالبیت به واریانس فنوتیپی (۰، ۰/۰۵ و ۰/۱۵) بر صحت انتخاب ژنومی در روش‌های آماری پارامتری (بیز A، B و LASSO) و ناپارامتری (RKHS) بررسی شد. همبستگی بین ارزش‌های اصلاحی حقیقی و ارزش‌های اصلاحی ژنومی، به عنوان معیاری از صحت پیش‌بینی‌های ژنومی برای سناریوهای مختلف به کمک نرم‌افزار R برآورد گردید. نتایج این مطالعه نشان داد که با افزایش مقدار وراثت‌پذیری در تمام روش‌ها صحت پیش‌بینی ژنومی افزایش پیدا می‌کند. همچنین با افزایش مقادیر نسبت واریانس غالبیت به واریانس فنوتیپی در مدل، شیب صحت ژنومی در روش‌های پارامتری کند گردیده ولی در روش ناپارامتری سرعت افزایش صحت ادامه پیدا نمود. میانگین مربعات خطا با افزایش مقدار نسبت واریانس غالبیت به واریانس فنوتیپی در روش ناپارامتری بیشتر کاهش یافت. از این رو می‌توان چنین نتیجه گرفت که با افزایش سطح نسبت واریانس غالبیت به واریانس فنوتیپی در مدل، استفاده از روش‌های ناپارامتری احتمالاً بهتر از روش‌های پارامتری صحت پیش‌بینی‌های ژنومی را برآورد می‌کنند.

واژه‌های کلیدی: اثر افزایشی، اثر غالبیت، انتخاب ژنومی، صحت انتخاب، شبیه‌سازی

مقدمه

امروزه در برنامه‌های اصلاح دام، حیوانات براساس ارزش اصلاحی پیش‌بینی شده انتخاب می‌شوند. پیشرفت ژنتیکی سریع نیازمند به پیش‌بینی‌های دقیق ارزش‌های اصلاحی به خصوص در اوایل زندگی حیوان است. این مهم توسط نشانگرهای مولکولی که به تعداد زیادی در سرتاسر ژنوم پراکنده شده‌اند، فراهم می‌گردد. انتخاب ژنومی که ارزش‌های اصلاحی ژنومی حیوان را اندازگیری می‌کند (GEBV)، حالتی از روش انتخاب به کمک نشانگر[۱] (MAS) می‌باشد. در این حالت اثر تعداد زیادی نشانگر که در سرتاسر ژنوم پراکنده شده‌اند، به طور همزمان برآورد می شود و برخلاف روش MAS که درصد کمی از واریانس ژنتیکی صفت را توجیه می‌کند، همه واریانس ژنتیکی به خصوص در صفاتی که اندازه‌گیری آنها مشکل، دارای وراثت‌پذیری کم، محدود به جنس و یا در اواخر عمر حیوان قابل اندازه‌گیری هستند، را توجیه می‌کنند (۱۸،۱۰).

پیش‌بینی دقیق ارزش اصلاحی ژنومی یکی از مباحث مهم و چالش‌برانگیز در اصلاح دام و نباتات است. صحت ارزش اصلاحی ژنومی تا حدودی به پارامترهای مدل و مدل‌های آماری که برای بدست آوردن اثرات نشانگری استفاده می‌شوند، ارتباط دارد. مطالعات شبیه‌سازی و تجربی نشان می‌دهد که صحت‌های قابل قبولی برای پیشرفت ژنتیکی در انتخاب ژنومی ایجاد شده است (۱۸،۹). با افزایش استفاده از روش انتخاب ژنومی، مدل‌ها و روش‌های آماری متعددی نیز برای پیش‌بینی ارزش اصلاحی ژنومی گسترش یافته است در نتیجه تعیین روش و مدل(های) آماری که بتواند به طور کارآمدتر و دقیق‌تر ارزش اصلاحی ژنومی را پیش‌بینی کند از اهمیت قابل توجهی به ویژه در زمان و هزینه برخوردار خواهد بود (۲۰،۱۹،۱۵).

یکی از عوامل تاثیرگزار بر صحت انتخاب ژنومی مدل‌های آماری می‌باشند. در این مدل‌ها به طور گسترده فقط اثرات ژنتیکی افزایشی وارد مدل گردیده است و تلاش‌های کمی در راستای قرار دادن اثرات ژنتیکی غیرافزایشی مثل اثر غالبیت یا اپیستازی در مدل صورت گرفته است. واریانس غالبیت به طورکلی خیلی کمتر از واریانس افزایشی می‌باشد (۲۴،۳). از طرف دیگر تخمین اثرات غیرافزایشی مشکل بوده است،زیرا مقدار این اثرات جزئی بوده و از طرف دیگر به طور رایج با اثرات مادری و محیط مشترک ترکیب می‌باشند (۲) به همین دلیل بیشتر سیستم‌های ارزیابی ژنتیکی از مدل‌های افزایشی استفاده نموده و از قرار دادن اثرات غیرافزایشی در مدل چشم‌پوشی می‌کنند، اما با این وجود باید گفته شود که قرار دادن اثرات غیرافزایشی (غالبیت و اپی‌ستازی) در مدل‌های پیش‌بینی ژنومی احتمالاً می‌تواند باعث افزایش صحت گردد.

در مطالعه‌ای با قرادادن اثرات غالبیت در مدل برای صفت رنگ پوشش بدن صحت انتخاب ژنومی را در موش ۱۷ درصد

1- Marker assisted selection

[Downloaded from rap.sanru.ac.ir on 2026-11-01]

[DOI: 10.29252/rap.8.18.161]

بیشتر برآورد نمودند (۱۷). در یک مطالعه گزارش گردید که ۱۰ درصد از تنوع ژنتیکی در ابتدای رشد تحت تاثیر اثر غالبیت قرار دارد (۶). در یک مطالعه با داده شبیه‌سازی شده نشان داده شد که با قرار دادن اثر غالبیت در مدل، نسبت به زمانی که این اثر وارد مدل نگردیده بود صحت انتخاب ژنومی افزایش یافت (۲۵). در مطالعه دیگر که روی گاوهای گوشتی برای صفات رشد، لاشه و باروری در روی داده واقعی صورت گرفت، گزارش داده شد که قرار دادن اثرات غالبیت در مدل تغییر معنی‌داری در صحت برآورد ارزش‌های اصلاحی ژنومی ایجاد نکرد و گزارش کردند که این موضوع احتمالاً به‌دلیل مقدار خیلی کم واریانس غالبیت نسبت به واریانس افزایشی و مشکل برآورد دقیق اثرات غالبیت باشد (۲). در مطالعه‌ای که روی مرغان تخم‌گذار انجام گردید، با قرار دادن اثراتغیرافزایشی (غالبیت) در مدل‌های ژنومی، صحت پیش‌بینی‌های ژنومی را با سه روش G-BLUP، P-BLUP و Bayes C مورد بررسی قرار دادند. نتایج آنها نشان داد که با وجود قرار دادن اثرات غالبیت در مدل، تغییری در افزایش صحت پیش‌بینی‌های ژنومی ایجاد نگردید. یک دلیل برای عدم بهبود در صحت با وجود قرار دادن اثرات غیرافزایشی در مدل را مشکل برآورد اثرات غالبیت به دلیل مقدار کم واریانس غالبیت بیان نمودند. همچنین گزارش نمودند که برای برآورد این اثرات باید از پنل‌های نشانگری با دانسیته بالا (بیشتر از ۵۰ کیلو باز) در مطالعات استفاده نمود (۱۴).گزارش شده است که روش‌های آماری ناپارامتری بهتر از روش‌های پارامتری (اگر مدل شامل اثرات غیرافزایشی باشد) ارزش‌های اصلاحی ژنومی را برآورد می‌کنند (۴). نشان داده شده که روش ناپارامتری کرنل برای داده جوجه‌های گوشتی نسبت به روش‌های پارامتری با صحت بالاتری ارزش‌های اصلاحی ژنومی را هنگامی که مدل شامل اثرات غیرافزایشی باشد، برآورد نموده است (۱۲). لذا هدف از مطالعه حاضر برآورد صحت ارزش‌های اصلاحی ژنومی با استفاده از داده شبیه‌سازی شده برای مدل‌های با اثرات افزایشی و سطوح مختلف اثرات غالبیت و مقایسه صحت انتخاب ژنومی به کمک روش‌های پارامتری و ناپارامتری با سطوح وراثت‌پذیری متفاوت است.

مواد و روش‌ها

ابتدا ۱۰۰۰ نسل مجزا با ۲۰۰۰ فرد شبیه‌سازی گردید، در ادامه برای ایجاد LD، از نسل ۱۰۰۰ به نسل ۲۰۰۰ تعداد جمعیت به ۱۰۰۰ نفر تقلیل پیدا نمود. تعداد افراد هر جنس در این مراحل با هم برابر بوده و سیستم تلاقی براساس جفت‌شدن تصادفی گامت‌هایی که نمونه‌های تصادفی از ژنگاه گامت‌های نر و ماده بودند، صورت گرفت (مرحله اول). سپس برای ایجاد یک جمعیت، تعداد ۵۰۰ حیوان از جمعیت پایه برای ۵ نسل تلاقی با افراد دیگر شبیه‌سازی شدند (مرحله دوم). به منظور افزایش اندازه جمعیت، ۸ نسل با تعداد ۱۰ نتاج برای هر ماده شبیه‌سازی گردید (مرحله سوم). تمام حیوانات نسل ۸ این مرحله به عنوان جمعیت مرجع برای تخمین اثرات نشانگر مورد بررسی قرار گرفتند. سپس برای یک نسل تلاقی، تعداد ۱۰۰ حیوان با ۱۰ نتاج برای هر ماده شبیه‌سازی گردیده و افراد این نسل برای جمعیت تأیید برای صحت پیش‌بینی ژنومی مورد استفاده قرار گرفتند. از نرم‌افزار QMSim (۲۲) برای ایجاد یک جمعیت پایه در طول زمان استفاده گردید. نسل‌های بعدی، انتخاب ژنومی و ارزیابی‌ها به کمک کدنویسی در نرم‌افزار R انجام شد.

ژنوم و فنوتیپ‌های صفات

ژنومی با یک کروموزوم یک مورگانی با ۱۰۰ QTL و ۱۰۰۰ نشانگر مجزا شبیه‌سازی گردید (جدول ۱). توزیع QTL و نشانگرها روی کروموزم به صورت تصادفی بود. در اوخر نسل پایه برای ایجاد یک پنل نشانگری، ۱۰۰۰ نشانگر که به صورت تفرق از نشانگرهای با دارا بودنMAF حداقل ۰/۰۵ بودند، ترسیم شد. اثرات افزایشی (a) برای یک QTL، برابر با نصف تفاوت ارزش ژنوتیپی بین هموزیگوت‌های جایگزین و اثرات غالبیت (d) بصورت انحراف ارزش هتروزیگوت از میانگین دو هموزیگوت تعریف می‌شود (۸). یک توزیع گاما با پارامترهای شکل و مقیاس به ترتیب ۰/۴ و ۱/۶۶ به منظور اثرات افزایشی برای هر QTL استفاده شد. وراثت‌پذیری در سه سطح ۰/۱، ۰/۳ و ۰/۵ و نسبت واریانس غالبیت به واریانس فنوتیپی در سه سطح ۰، ۰/۰۵ و ۰/۱۵ در نظر گرفته شد.

جدول ۱- ساختار جمعیت پایه و پارامترهای شبیه‌سازی شده

Table 1. The structure of base populations and simulation parameters

پارامتر	مقدار	پارامتر	مقدار
ایجاد عدم تعادل پیوستگی	نسل ۱ تا ۲۰۰۰	تعداد نشانگرها در طول ژنوم	۱۰۰۰
جمعیت برای افزایش تعداد	۲۰۰۰ تا ۲۰۵۵	تعداد جایگاه صفات کمی	۱۰۰
جمعیت مرجع	۲۰۵۵ تا ۲۰۶۳	نرخ جهش	$۲/۵\times۱۰^{-۴}$
جمعیت هدف	۲۰۶۳ تا ۲۰۶۴	طول ژنوم (cM)	۱
تعداد کروموزوم	۱	اثرات افزایشی آل‌ها برای QTL	گاما
تعداد جمعیت مرجع	۱۰۰۰		

تخمین اثرات نشانگر

برای تخمین اثرات نشانگر از روش‌های پارامتری (BayesA، BayesB وBayes LASSO) و روش ناپارامتری RKHS استفاده شد. تفاوت در روش‌های پارامتری در توزیع پیشین اثرات نشانگر می‌باشد. مدل آماری برای تخمین اثرات نشانگر در روش‌های پارامتری به‌صورت معادله ۱ می‌باشد:

$$y_i = \mu + \Sigma X_{ij} a_j + \Sigma Z_{ij} d_j + e \quad \text{معادله (۱)}$$

در معادله ۱، y_i: ارزش فنوتیپی فرد i در جمعیت مرجع، μ:

[Downloaded from rap.sanru.ac.ir on 2026-11-01]

[DOI: 10.29252/rap.8.18.161]

میانگین کل، X_{ij} : تعداد کپی الل‌های نشانگر j، که به ترتیب کد ۰، ۱ و ۲ برای aa، Aa وAA می‌باشد، a_j: اثر ناشناخته تصادفی افزایشی برای نشانگرj، Z_{ij}: متغیر شاخص برای هتروزیگوسیتی فرد i در نشانگرj، $Z_{ij}=0$ وقتیکه فرد i در مارکر j هموزیگوس (AA یا aa) و $Z_{ij}=1$ وقتیکه فرد i در مارکر j هتروزیگوس (Aa) باشد. d_j: اثر غالبیت ناشناخته تصادفی برای نشانگرj، $\mathbf{e}_i$: اثرات تصادفی برای حیوان i و Σ: مجموع تمام نشانگرهای لوکاس j می‌باشد. توزیع پیشین برای واریانس باقی‌مانده برابر با $\sigma_e^2 \sim x^{-2}(df_e, S_e)$ بود که درجه آزادی و پارامتر مقیاس (S_e) به ترتیب ۳/۵ و ۳ بودند. توزیع پیشین شرطی برای اثرات نشانگر از توزیع نرمال با واریانس پیشین برای هر نشانگر به‌صورت $a_{j \sim N(0,\sigma^2 \in t^2)}$ تعریف می‌گردد که j=1,....,m و t_j^2 با توزیع پیشین نمایی به‌صورت $t_j^2 = \exp(a^2)$ تعریف می‌گردد. توزیع پیشین برای واریانس غالبیت برابر با $d^2 \sim i^{-2}(df_d, S_d)$ بود، درجه آزادی و پارامتر غالبیت (S_d) به ترتیب ۳ و ۰/۰۰۰۵ در نظر گرفته شد. از بسته نرم‌افزاری BLR (۲۱) در نرم‌افزار R برای محاسبه توزیع پیشین پارامترها استفاده گردید. برای محاسبه اثرات نشانگر در تمام مدل‌های بیزی از الگوریتم MCMC استفاده گردید. بدین منظور از یک زنجیره با طول ۱۰۰۰۰ تکرار استفاده گردید. ۱۵۰۰ نمونه اول به عنوان دوره گرم کردن استفاده شد.

فضای هیلبرت با هسته‌ی بازآفرین
ReproducingKernel Hilbert Spaces (RKHS)

این روش به عنوان یک روش ناپارامتری برای برآورد ارزش اصلاحی ژنومی معرفی شده است (۱۱). تخمین پارامترهای مدل به کمک مجموع مربعات جریمه انجام می‌گیرد. مدل آماری این روش به صورت معادله ۲ می‌باشد:

معادله (۲)

$$y_i = w_i' + z'u + g(x_i) + e_i,$$

در معادله ۲، β_i،i=1,2,...,n بردار اثرات نامعین ثابت، **u** یک بردار q×1 از اثرات ژنتیکی افزایشی، w_i' و z_i' بردارهای مشخص ضرایب، $g(x_i)$ و $\mathbf{e}_i$ به ترتیب تابع نامشخص از داده‌های نشانگر و بردار باقیمانده و **u** توزیع آن به صورت $N(0, A\pm_u^2)$ که A ماتریس خویشاوندی افزایشی می‌باشد. مجموع مربعات جریمه به صورت زیر می‌باشد:

معادله (۳)

$$SS(g(x),h) = \sum_{i=1}^{n}[y_i - w_i' - z_i'u - g(x_i)]^2 + h\|g(x)\|,$$

در معادله ۳، مقدار جریمه $\|g(x)\|$ یک تابع از مشتق دوم g(x) است، هدف از مشتق دوم به حداقل رساندن مقدار جریمه SS می‌باشد. RKHS به کمک روش‌ها و الگوریتم‌های توصیف شده (۷) و به کمک نرم‌افزار R صورت گرفت.

صحت پیش‌بینی ژنومی

در این مطالعه توانایی پیش‌بینی مدل‌ها از سه مسیر مورد بررسی قرار گرفت: (۱) همبستگی بین ارزش‌های اصلاحی ژنومی و ارزش‌های اصلاحی واقعی در جمعیت تأیید (۲) میانگین مربعات خطای پیش‌بینی ژنومی (MSE) که از رابطه زیر پیش‌بینی گردید (۱۶):

معادله (۴)

$$MSE = \frac{1}{N}\sum (GBV - TBV)^2$$

در معادله ۴، GBV و TBV به ترتیب ارزش اصلاحی ژنومی و ارزش اصلاحی واقعی بود. (۳) ضریب رگرسیون پیش‌بینی ژنومی،که از ضریب رگرسیون ارزش اصلاحی واقعی بر ارزش اصلاحی ژنومی بدست آمد، انتظار این‌است که این ضریب نزدیک به یک باشد که در این حالت نشان‌دهنده این‌است که مقیاس ارزش اصلاحی ژنومی مشابه ارزش اصلاحی تخمینی می‌باشد و اریب آن کمتر می‌باشد.

نتایج و بحث

صحت پیش‌بینی ارزش‌های اصلاحی ژنومی برای سه روش پارامتری و یک روش ناپارامتری با وراثت‌پذیری‌ها و نسبت واریانس‌های غالبیت متفاوت در جدول ۲ ارائه شده است. در تمام روش‌ها، با افزایش وراثت‌پذیری از ۰/۱ به ۰/۵ صحت پیش‌بینی‌های ژنومی افزایش پیدا نمود. دامنه صحت پیش‌بینی ژنومی از ۰/۶۱۲ برای روش RKHS تا ۰/۸۱۰ برای روش Bayes B متفاوت بدست آمد. نشان داده شده است که برای صفات با وراثت‌پذیری پایین، برای رسیدن به صحت ژنومی مورد قبول باید تعداد رکوردها را برای تخمین اثرات نشانگر افزایش داد (۵ و ۱۳). در مطالعه‌ای، با افزایش سطح وراثت‌پذیری از ۰/۱ به ۰/۵ صحت پیش‌بینی‌های ژنومی از ۰/۴۸ به ۰/۶۷ افزایش پیدا نمود (۴) همچنین در مطالعه دیگر نیز با افزایش میزان وراثت‌پذیری صحت پیش‌بینی ژنومی از ۰/۴۰ به ۰/۷۰ افزایش نمود (۲۳) که این نتایج با نتایج مطالعه حاضر مطابقت دارد.

از طرف دیگر با افزایش سهم نسبت واریانس غالبیت در واریانس فنوتیپی، صحت تخمین ژنومی در روش‌های پارامتری و ناپارامتری افزایش پیدا نمود. همانطور که در جدول ۲ مشاهده می‌گردد با افزایش سطح نسبت واریانس غالبیت از صفر به ۰/۰۵ برای تمام روش‌ها صحت تخمین ژنومی افزایش پیدا نمود، اما با افزایش سطح نسبت واریانس غالبیت از ۰/۰۵ به ۰/۱۵ شیب افزایش صحت تخمین ژنومی در روش‌های پارامتری کند گردید، ولی برای روش ناپارامتری افزایش صحت ژنومی ادامه پیدا نمود. در مطالعه‌ای که با داده شبیه‌سازی انجام گردید، گزارش داده شد که اگر معماری ژنتیکی صفات فقط براساس واریانس افزایشی باشد، روش‌های پارامتری بهتر از روش‌های ناپارامتری صحت ژنومی را نشان دادند. همچنین، نشان دادند که با افزایش سهم واریانس غالبیت در مدل‌های ژنومی، صحت تخمین ژنومی در روش‌های ناپارامتری بیشتر از روش‌های پارامتری گزارش نمودند، که نتایج آنها در تأیید مطالعه حاضر بود (۱۶). نتایج این مطالعه نشان داد که اگر تنها اثرات افزایشی در مدل شبیه‌سازی شوند، مدل‌های پارامتری نسبت به مدل‌های

[Downloaded from rap.sanru.ac.ir on 2026-11-01]
[DOI: 10.29252/rap.8.18.161]

ناپارامتری توان صحت پیش‌بینی ژنومی آنها بیشتر بوده زیرا فرض بر این است که نشانگرها مستقل از هم هستند، اما در حالت شبیه‌سازی اثرات غالبیت علاوه بر اثرات افزایشی در مدل، با توجه به وابسته بودن نشانگرها و نقض فرض مدل‌های پارامتری، روش‌های ناپارامتری بهتر از روش‌های پارامتری توان پیش‌بینی را دارا هستند.

متوسط صحت تخمین ارزش‌های اصلاحی ژنومی برای تمام سناریوهای مختلف برای روش‌های پارامتری بیز A، B و LASSO به ترتیب ۰/۶۹۸، ۰/۷۴۹ و ۰/۷۰۶ بدست آمد، که صحت تخمین ژنومی در روش بیز B بیشتر از دو روش دیگر می‌باشد. در مطالعه‌ای برای بررسی صحت تخمین ارزش اصلاحی ژنومی به کمک داده‌های شبیه‌سازی شده که به کمک روش‌های حداقل مربعات،BLUP، بیز A و B صورت گرفت صحت تخمین ژنومی در روش بیز B نسبت به بقیه روش‌ها بیشتر بدست آمد که با مطالعه حاضر مطابقت داشت (۱۸).

اگرچه وراثت‌پذیری توانایی تشخیص اینکه کدام روش انتخاب ژنومی صحت آن نسبت به دیگر روش‌ها بیشتر است را ندارد، اما مشاهده گردید که وقتی وراثت‌پذیری بالا و معماری ژنتیکی افزایشی بود، صحت تخمین ژنومی نسبت به وقتی که وراثت‌پذیری پایین می‌باشد، بیشتر است (جدول ۲).

جدول ۲- صحت تخمین ارزش اصلاحی ژنومی با روش‌های پارامتری و نا پارامتری برای وراثت‌پذیری‌ها و نسبت‌های واریانس غالبیت متفاوت

Table 2. The accuracy of the prediction of genomic breeding value for parametric and non-parametric methods with heritability and different dominance variance.

وراثت‌پذیری	واریانس غالبیت	BayesA	BayesB	BayesLASSO	RKHS
	۰	۰/۶۳۳	۰/۶۹۷	۰/۶۵۴	۰/۶۱۲
۰/۱	۰/۰۵	۰/۶۹۸	۰/۷۳۳	۰/۶۹۷	۰/۶۷۳
	۰/۱۵	۰/۶۴۵	۰/۷۲۲	۰/۶۷۰	۰/۶۹۵
	۰	۰/۶۸۱	۰/۷۴۲	۰/۶۹۶	۰/۶۳۲
۰/۳	۰/۰۵	۰/۷۲۳	۰/۷۶۸	۰/۷۲۲	۰/۶۷۴
	۰/۱۵	۰/۷۰۹	۰/۷۲۱	۰/۷۱۹	۰/۷۰۲
	۰	۰/۷۳۲	۰/۷۸۲	۰/۷۹۰	۰/۶۸۹
۰/۵	۰/۰۵	۰/۷۴۱	۰/۸۱۰	۰/۷۱۲	۰/۷۱۲
	۰/۱۵	۰/۷۲۳	۰/۷۶۸	۰/۷۰۲	۰/۷۲۱

متوسط میانگین مربعات خطا در روش‌های بیز A، B و LASSO و روش RKHS به ترتیب ۱/۸، ۱/۷، ۱/۸۶ و ۱/۹۵ برآورد گردید (جدول ۳). همانطور که درجدول ۳ مشاهده می‌شود با افزایش وراثت‌پذیری از ۰/۱ به ۰/۵ میانگین مربعات خطا برای روش‌های پارامتری کاهش پیدا کرد، ولی با افزایش سهم اثرات غیرافزایشی (غالبیت) در مدل برای روش‌های پارامتری میانگین مربعات خطا افزایش پیدا نمود. در مقابل با افزایش مقدار وراثت‌پذیری میانگین مربعات خطا در روش ناپارامتری افزایش پیدا نمود، ولی با افزایش سهم اثرات غیرافزایشی غالبیت در مدل‌های ژنومی میانگین مربعات خطا در روش ناپارامتری RKHS کاهش پیدا نمود.

جدول ۳- میانگین مربعات خطای پیش‌بینی ارزش اصلاحی ژنومی با روش‌های پارامتری و ناپارامتری برای وراثت‌پذیری‌ها و نسبت‌های واریانس غالبیت متفاوت

Table 3. The mean squared prediction error of genomic breeding value for parametric and non parametric metl The with heritability and different dominance variance

وراثت‌پذیری	واریانس غالبیت	BayesA	BayesB	BayesLASSO	RKHS
	۰	۱/۹۳	۱/۸۳	۱/۹۲	۱/۹۸
۰/۱	۰/۰۵	۱/۷۵	۱/۶۵	۱/۷۷	۱/۸۳
	۰/۱۵	۱/۸۳	۱/۷۳	۱/۸۴	۱/۷۳
	۰	۱/۸۵	۱/۷۵	۱/۸۹	۲/۱۱
۰/۳	۰/۰۵	۱/۷۷	۱/۶۷	۱/۹۳	۱/۹۳
	۰/۱۵	۱/۸۲	۱/۷۲	۲/۱۱	۱/۸۷
	۰	۱/۶۹	۱/۵۹	۱/۷۶	۲/۲۱
۰/۵	۰/۰۵	۱/۷۶	۱/۶۶	۱/۷۲	۱/۹۸
	۰/۱۵	۱/۸۱	۱/۷۱	۱/۸۳	۱/۸۸

در یک مطالعه که با داده شبیه‌سازی شده صورت گرفت، با افزایش وراثت‌پذیری میانگین مربعات خطا در روش‌های پارامتری کاهش و در روش‌های ناپارامتری افزایش پیدا نمود، همچنین گزارش داده شد که با افزایش سهم اثرات غیرافزایشی در مدل، میانگین مربعات خطا در روش‌های پارامتری افزایش و در روش‌های ناپارامتری کاهش پیدا نمود (۱۶). همچنین در مطالعه‌ای که با داده واقعی صورت گرفت، MSE با مدل داری اثر غالبیت نسبت به مدلی که این اثر را نداشت کاهش بیشتری داشت (۱) که با نتایج مطالعه حاضر مطابقت دارد. این نتایج نشان می‌دهد که شایستگی برازش مدل و متعاقب آن صحت پیش‌بینی ژنومی می‌تواند به وسیله وارد کردن اثرات غالبیت در مدل‌های ژنومی برای پیش‌بینی عملکرد حیوانات در آینده، بهبود پیدا کند. یکی از پیامدهای مهم نادیده گرفتن اثرات غیرافزایشی در مدل‌های ژنومی،

[Downloaded from rap.sanru.ac.ir on 2026-11-01]

[DOI: 10.29252/rap.8.18.161]

ارزیابی‌های افراد و متعاقباً طبقه‌بندی نادرست افراد براساس ارزش‌های اصلاحی تخمینی باشد. در مطالعه‌ای با وارد کردن اثرات غالبیت در مدل‌های ارزیابی ژنومی صحت تخمین واریانس ژنتیکی افزایشی و ارزش‌های اصلاحی افراد بالاتر رفته، بنابراین منتهی به صحت بیشتر پیش‌بینی‌های ژنومی گردیده است (۱۷).

میانگین ضریب رگرسیون پیش‌بینی ژنومی در روش‌های بیز A، B، LASSO و RKHS بترتیب ۰/۸۰، ۰/۹۲، ۰/۸۱ و ۰/۸۴ بدست آمد (جدول ۴). با توجه به اعداد بدست آمده، روش بیز B نسبت به دیگر روش‌ها ناریب‌تر بود. در روش‌های پارامتری (بیز A، B و LASSO) با افزایش وراثت‌پذیری از ۰/۱ به ۰/۵ ضریب رگرسیون به عدد یک نزدیکتر و اریب آنها کمتر می‌گردد، از طرف دیگر با وارد کردن اثرات غیرافزایشی غالبیت در مدل اریب این روش‌ها بیشتر گردید (اعداد از یک دور می‌گردند). در روش ناپارامتری RKHS با افزایش وراثت‌پذیری، مقدار اریب بیشتر شده و با وارد کردن اثرات غیرافزایشی غالبیت در مدل، ضریب رگرسیون پیش‌بینی ژنومی به عدد یک نزدیکتر بنابراین اریب کاهش پیدا نمود. در یک مطالعه که برای صفات باروری و تولید شیر در گاوهای شیری انجام گردید، میانگین ضریب رگرسیون فنوتیپ تصحیح شده بر ارزش اصلاحی تخمینی برای مدلی که شامل هر دو اثر بود نسبت به مدلی که فقط دارای اثر افزایشی بود، به عدد یک نزدیکتر و لذا توانایی پیش‌بینی بهتر مدل دارای اثر غالبیت را روشن می‌کند. که در توافق با مطالعه حاضر بود (۱).

جدول ۴- ضریب رگرسیون پیش‌بینی ژنومی با روش‌های پارامتری و نیمه‌پارامتری برای وراثت‌پذیری‌ها و نسبت واریانس غالبیت متفاوت

Table 4. The regression coefficient prediction of genomic breeding value for parametric and non-parametric methods with heritability and different dominance variance

وراثت‌پذیری	واریانس غالبیت	BayesA	BayesB	BayesLASSO	RKHS
	۰	۰/۸۳	۰/۹۴	۰/۸۴	۰/۸۱
۰/۱	۰/۰۵	۰/۷۹	۰/۹۲	۰/۷۹	۰/۸۳
	۰/۱۵	۰/۷۸	۰/۸۹	۰/۷۸	۰/۹۰
	۰	۰/۸۴	۰/۹۵	۰/۸۸	۰/۷۹
۰/۳	۰/۰۵	۰/۸۰	۰/۹۲	۰/۷۹	۰/۸۲
	۰/۱۵	۰/۷۲	۰/۹۰	۰/۷۸	۰/۹۲
	۰	۰/۸۶	۰/۹۷	۰/۹۲	۰/۷۸
۰/۵	۰/۰۵	۰/۸۰	۰/۹۱	۰/۸۸	۰/۸۲
	۰/۱۵	۰/۷۴	۰/۹۰	۰/۸۳	۰/۹۳

بطور کلی نتایج این مطالعه نشان داد که ۱- با افزایش سطوح وراثت‌پذیری، صحت پیش‌بینی ژنومی در روش‌های پارامتری و ناپارامتری افزایش پیدا نمود. ۲- با افزایش سطح غالبیت در مدل، صحت پیش‌بینی ژنومی برای هر دو روش آماری افزایش پیدا نمود، اما در نسبتی ازسطح غالبیت برای روش‌های پارامتری صحت کند گشته ولی در روش ناپارامتری ادامه افزایش صحت دیده شد. ۳- MSE و شیب رگرسیون برآورد شده نشان داد که مدل‌های دارای اثر غالبیت، نسبت به مدل‌های که این اثر را نداشتند دارای برازش بهتر و اریب کمتری بودند. به طور کلی با افزایش سطح اثرات ژنتیکی غیرافزایشی (غالبیت) در مدل‌های ژنومی استفاده از روش‌های ناپارامتری نسبت به روش‌های پارامتری برای برآورد صحت پیش‌بینی‌های ژنومی توصیه می‌گردد.

[Downloaded from rap.sanru.ac.ir on 2026-11-01]

[DOI: 10.29252/rap.8.18.161]

منابع

1. Aliloo, H., J.E. Pryce, O. Gonzalez-Recio, B.G. Cocks and B.J. Hayes. 2016. Accounting for dominance to improve genomic evaluations of dairy cows for fertility and milk production traits. Genetic Selection Evolution, 48: 8-11.
2. Bolormaa, S., J.E. Pryce1, Y. Zhang, A. Reverter, W. Barendse, B.J. Hayes and M.E. Goddard. 2015. Non-additive genetic variation in growth, carcass and fertility traits of beef cattle. Genetic Selection Evolution, 47(26): 12 pp.
3. Boysen, T.J., C. Heuer, J. Tetens, F. Reinhardt and G. Thaller. 2013. Novel use of derived genotype probabilities to discover significant dominance effects for milk production traits in dairy cattle. Genetics, 193: 431-42.
4. Calus, M.P.L. 2010. Genomic breeding value prediction: methods and procedures. Animal, 4(02): 157-164.
5. Calus, M.P.L., T.H.E. Meuwissen, A.P.W. De Roos and R.F. Veerkamp. 2008. Accuracy of genomic selection using different methods to define haplotypes. Genetics, 178: 553-561.
6. Carlborg, Ö., S. Kerje, K. Schütz,L. Jacobsson, P. Jensen and L. A. Andersson.2003.Global search reveals epistatic interaction between QTL for early growth in the chicken. Genome Research**,** 13: 413-421.
7. de los Campos, G., D. Gianola, G.J.M. Rosa, K.A. Weigel and J. Crossa. 2010. Semi-parametric genomic-enabled prediction of genetic values using reproducing kernel Hilbert spaces methods. Genetic Research, 92: 295-308.
8. Falconer, D.S. and T.F.C. Mackay. 1996. Introduction to Quantitative Genetics. 4th[ed]. Harlow: Pearson Education Limited.
9. Gao, H., O.F. Christensen, P. Madsen, U.S. Nielsen, Y. Zhang, M.S. Lund and G. Su. 2012. Comparison on genomic predictions using three GBLUP methods and two single-step blending methods in the Nordic Holstein population. Genetic Selection Evolution, 44(8): 8 pp.
10. Gianola, D. and J.B. Van Kaam. 2008. Reproducing Kernel Hilbert Spaces Regression methods for genomic assisted prediction of quantitative traits. Genetics, 178: 2289-2303.
11. Goddard, M.E. and B.J. Hayes. 2007. Genomic selection. Journal Animal Breeding and Genetics, 124: 323-330.
12. González-Recio, O., D. Gianola, N. Long, K.A. Weigel, G.J.M. Rosa and S. Avendaño. 2008. Nonparametric methods for incorporating genomic information into genetic evaluations: An Application to Mortality in Broilers. Genetics, 178: 2305-2313.
13. Hayes, B.J., P.J. Bowman, A.J. Chamberlain and M.E. Goddard. 2009. Invited review: Genomic selection in dairy cattle: Progress and challenges. Journal Dairy Science, 92: 433-443.
14. Heidaritabar, M., A. Wolc, J. Arango, J. Zeng, P. Settar, J.E. Fulton, N.P. Osullivan, J.W.M. Bastiaansen, R.L. Fernando, D.J. Garrick and J.C.M. Dekkers. 2016. Impact of fitting dominance and additive effects on accuracy of genomic prediction of breeding values in layers. Journal Animal Breeding and Genetics, 1-13.
15. Heslot, N., H.P. Yang, M.E. Sorrells and J.L. Jannink. 2012. Genomic selection in plant breeding: A Comparison of Models. Crop Science, 52: 146-160.
16. Howard, R., A.L. Carriquiry and W.D. Beavis. 2014. Parametric and nonparametric statistical methods for genomic selection of traits with additive and epistatic genetic architectures. G3, 4: 1027-1046.
17. Lee, S. H., J.H.J. Vander Werf, B.J. Hayes, M.E. Goddard and P.M. Visscher. 2008. Predicting unobserved phenotypes for complex traits from whole-genome SNP data. PLoS Genetic, 4: e1000231.
18. Meuwissen, T.H., B.J. Hayes and M.E. Goddard. 2001. Prediction of total genetic value using genome-wide dense marker maps. Genetics, 157: 1819-1829.
19. Neves, H., R. Carvalheiro and S.A. Queiroz. 2012. A comparison of statistical methods for genomic selection in a mice population. BMC Genetics, 13:100. 17 pp.
20. Ogutu, J.O., T. Schulz-Streeck and H.P. Piepho. 2012. Genomic selection using regularized linear regression models: ridge regression, lasso, elastic net and their extensions. BMC Proceedings. 6(Suppl 2): S10.
21. Perez, P., G. de los Campos, J. Crossa and D. Gianola. 2010. Genomic-enabled prediction based on molecular markers and pedigree using the Bayesian linear regression package in R. Plant Genome, 3: 106-16.
22. Sargolzaei, M. and F.S. Schenkel. 2009. QMSim: a large-scale genome simulator for livestock. Bioinformatics, 25: 680-1.
23. Saatchi, M., S.R. Miraei- Ashtiani, A. Nejati- Javaremi, M. Moradi-Shahrebabak and H. Mehrabani-Yeghaneh. 2010. The impact of information quantity and strength of relationship between training set and validation set on accuracy of genomic estimated breeding values. African Journal of Biotechnology, 9: 438-442.
24. Su, G., O.F. Christensen, T. Ostersen, M. Henryon and M.S. Lund. 2012. Estimating additive and non-additive genetic variances and predicting genetic merits using genome-wide dense single nucleotide polymorphism markers. PLoS One, 7(45): 293.
25. Wittenburg, D., N. Melzer and N. Reinsch. 2011. Including non-additive genetic effects in Bayesian methods for the prediction of genetic values based on genome-wide markers. BMC Genetics, 12-74.

[Downloaded from rap.sanru.ac.ir on 2026-11-01]

[DOI: 10.29252/rap.8.18.161]

Genomic Selection Accuracy Parametric and Nonparametric Statistical Methods with Additive and Dominance Genetic Architectures

Yahya Mohammadi[1] and Morteza Sattaei Mokhtari[2]

1- Assistant Professor, Department of Animal Science,University of Ilam,
(Corresponding Author: mohamadi_yahya@yahoo.com)
2- Assistant Professor, Department of Animal Science, University of Jiroft
Received: February 13, 2017 Accepted: Jun21, 2017

Abstract

In most genomic prediction studies only additive effects will be used in models for estimating genomic breeding values (GEBV). However, dominance genetic effects are an important source of variation for complex traits, considering them into account may improve the accuracy of GEBV. In the present study, performed applying simulated data, the effect of different heritability values (0.1, 0.3 and 0.5) and different values for the proportion of dominance variance to phenotypic variance (0, 0.05 and 0.15) on genomic selection accuracy in parametric (LASSO, A and B Bayes) and non-parametric (RKHS) statistical methods were studied. Correlations between the true and genomic breeding values, as a measure for the accuracy of genomic predictions under different scenarios were calculated using R software. The results of the present study revealed that, under all statistical methods as heritability values increased, the accuracy of genomic predictability increased. Also, as the value of dominance variance to phenotype variance increased genomic accuracy slant was slow under parametric methods but in the non-parametric method accuracy continued to increase. Under non parametric method the average mean square error was more reduced as the ratio of dominance variance to phenotype variance increased. Therefore, it may be concluded that under non-parametric method as the ratio of dominance variance to phenotype variance increased the accuracy of genomic predictions would be more increased than that of under parametric methods.

Keywords: Additive effect, Dominance effect, Genomic selection, Selection accuracy, Simulation

[Downloaded from rap.sanru.ac.ir on 2026-11-01]

[DOI: 10.29252/rap.8.18.161]