

کاربرد تجزیه‌ی مؤلفه‌های اصلی در ارزیابی ژنتیکی گاوهای شیری ایران

اعظم اسمعیلی نوجه‌ده^۱، صادق علیجانی^{۲*}، کریم حسن پور^۳ و آرش جوانمرد^۳

تاریخ دریافت: ۹۶/۳/۲۲ تاریخ پذیرش: ۹۶/۸/۲۹

^۱ دانشجوی کارشناسی ارشد گروه علوم دامی دانشکده کشاورزی دانشگاه تبریز

^۲ دانشیار گروه علوم دامی دانشگاه تبریز

^۳ استادیار گروه علوم دامی دانشگاه تبریز

* مسئول مکاتبه: Email: sad-ali@tabrizu.ac.ir

چکیده

زمینه مطالعاتی: مدل رگرسیون تصادفی یکی از دقیق‌ترین مدل‌ها برای پیش‌بینی ارزش اصلاحی، با استفاده از رکوردهای روزآزمون می‌باشد. با این حال به‌کار بردن این مدل از نظر محاسباتی دشوار و زمان‌بر است. هدف: تعیین اهمیت نسبی ارزش‌های اصلاحی در روزهای مختلف شیردهی و برآورد مؤلفه‌های اصلی ژنتیکی برای ارزش‌های اصلاحی صفات تولید شیر گاوهای هلشتاین ایران اهداف اصلی تحقیق حاضر می‌باشند. روش کار: از رکوردهای روزآزمون تولید شیر، درصد چربی و درصد پروتئین دوره شیردهی اول گاوهای شیری هلشتاین (متولد سال‌های ۱۳۶۷ تا ۱۳۹۴) که توسط مرکز اصلاح نژاد کشور جمع‌آوری شده بود، استفاده شد. برای صفات تولید شیر، درصد چربی و درصد پروتئین به‌ترتیب از رکوردهای ۷۳۸۳۹، ۶۵۱۶۵ و ۴۶۸۸۱ رأس گاو، از ۲۳۰ گله که در شجره خود دارای ۱۷۶۳۹۰ رأس گاو بود، استفاده شد. پارامترهای ژنتیکی این صفات با استفاده از مدل رگرسیون تصادفی و توسط GIBSS3F90 برآورد شد. سپس ماتریس همبستگی بین ارزش‌های اصلاحی به‌دست‌آمده در روزهای مختلف شیردهی محاسبه گردید. در ادامه، مؤلفه‌های اصلی ژنتیکی از ارزش‌های اصلاحی توسط رویه PRINCOMP نرم‌افزار SAS به‌دست آمد. **نتایج:** ماتریس همبستگی ژنتیکی بین ارزش‌های اصلاحی پیش‌بینی‌شده در روزهای مختلف نشان می‌دهد که ارزش‌های اصلاحی در اواسط دوره شیردهی برای تمامی صفات همبستگی بالایی دارند. با استفاده از تجزیه‌ی مؤلفه‌های اصلی برای ارزش‌های اصلاحی مشاهده شد که دو مؤلفه‌ی اصلی اول درصد بالایی از واریانس ژنتیکی کل را تبیین می‌کنند. برای صفت تولید شیر اولین مؤلفه‌ی اصلی ۹۹/۴۸ درصد و برای صفات درصد چربی و درصد پروتئین به ترتیب ۹۸/۱۹ درصد و ۱۰۰ درصد از واریانس کل ژنتیکی توسط دو مؤلفه اصلی اول تبیین شد. **نتیجه‌گیری نهایی:** در جهت کاهش هزینه‌های رکوردبرداری و با در نظر گرفتن همبستگی بالای بین ارزش‌های اصلاحی به‌نظر می‌رسد، پیش‌بینی ارزش‌های اصلاحی برای کل روزهای آزمون ضرورتی ندارد. بنابراین می‌توان روی رکوردبرداری در روزهایی که با مؤلفه‌های اصلی ارتباط بالایی نشان می‌دهند، تمرکز نمود.

واژگان کلیدی: رگرسیون تصادفی، ژنتیک، مؤلفه اصلی، همبستگی ژنتیکی

مقدمه

مدل رگرسیون تصادفی برای تجزیه و تحلیل ژنتیکی صفات، جزء یکی از انتخاب‌های مطلوب محسوب می‌شود (خیرآبادی و همکاران ۱۳۹۲) که در طول زمان و به تدریج دستخوش تغییرات شد و توسط افراد مختلف به کار گرفته شد (عزیز و همکاران ۲۰۰۵). یکی از مزیت‌های مدل رگرسیون تصادفی این است که در آن واریانس‌های مختلف نظیر ژنتیکی و محیط دائمی، در طول دوره رکوردگیری ثابت در نظر گرفته نمی‌شود و برای هر دوره رکوردگیری اثر محیطی خاصی منظور می‌شود (کرانیس و همکاران ۲۰۰۷). مدل رگرسیون تصادفی قادر به پیش‌بینی ساختار کواریانس بین رکوردهای روز آزمون (TDR) در طول یک مقیاس زمانی پیوسته می‌باشد و برخلاف مدل چند صفتی، اجازه پیش‌بینی BV برای کل یا بخشی از دوره شیردهی را می‌دهد (تاکما و آکباش ۲۰۰۹ و بیگناردی و همکاران ۲۰۰۹). مزایای دیگر استفاده از مدل رگرسیون تصادفی شامل بهبود صحت انتخاب، نیاز به پارامترهای کم برای توصیف صفات نسبت به مدل‌های چند صفتی و عدم نیاز به استانداردسازی سن می‌باشد (لیوایز و پروترستون ۲۰۰۲ و شفر ۲۰۰۴). استفاده از مدل رگرسیون تصادفی سبب می‌گردد تا انعطاف‌پذیری بیشتری در مدل‌سازی آماری بدون تعریف آستانه خاصی صورت گیرد (بهلولی و همکاران ۲۰۱۳).

محمدی و همکاران (۱۳۹۱) نشان دادند که با توجه به معیارهای مقایسه، مناسب‌ترین تابع در مدل رگرسیون تصادفی، جهت برآورد پارامترهای ژنتیکی صفات تولید شیر در گاوهای هلشتاین ایران، چندجمله‌ای لژاندر، می‌باشد؛ با این حال یکی از مشکلات اساسی مدل رگرسیون تصادفی با توجه به مطالعات پژوهشی گسترده انجام شده، بالا بودن حجم محاسباتی می‌باشد و گاه‌ها تا ماه‌ها این محاسبات به طول می‌انجامد (جنسن ۲۰۰۱). بنابراین، کاهش ابعاد محاسبه در رگرسیون تصادفی برای کاهش هزینه‌های محاسباتی، خیلی مهم می‌باشد. به این خاطر محققان درصدد کاهش این

اطلاعات می‌باشند. شایع‌ترین این روش‌ها استفاده از تجزیه و تحلیل مؤلفه‌های اصلی است.

تجزیه و تحلیل مؤلفه‌های اصلی (PCA) یک روش ریاضی می‌باشد که با استفاده از یک تبدیل متعامد برای کاهش مجموعه متغیرهای همبسته به مجموعه متغیرهای ناهمبسته به نام مؤلفه اصلی (PC) می‌باشد (ساوگناگو و همکاران ۲۰۱۱). به عبارت دیگر PCA یک تبدیل در فضای برداری است که بیشتر برای کاهش ابعاد مجموعه داده‌ها مورد استفاده قرار می‌گیرد. PC را می‌توان در شاخص انتخاب مورد استفاده قرارداد که باعث ساده شدن آن می‌شود، در چنین شاخصی می‌توان از چند PC به جای تمام صفات اصلی استفاده کرد. مزیت اصلی این مؤلفه‌ها استقلال آن‌هاست. از نظر پیشرفت ژنتیک حیوانی، گروهی از PC های ویژه‌ای که به‌طور هم‌زمان ممکن است برای اهداف انتخاب جالب باشند در نظر گرفته می‌شوند. جنبه مهم دیگر این است که هر یک از این PC ها درصدی از واریانس کل را تشکیل می‌دهند. PCA یک ابزار مناسب برای ارزیابی و درک واریانس کل نشأت گرفته از گروهی از صفات وابسته به هم می‌باشد، این کار اجازه کاهش تعداد صفات را می‌دهد که در شاخص انتخاب برنامه‌های اصلاح نژاد در نظر گرفته می‌شوند (باتیستا پنتو و همکاران ۲۰۰۶).

برای برآورد ارزش‌های اصلاحی (PBV) روزانه از رکوردهای ماهانه انفرادی کل گاوها در طول دوره شیردهی استفاده شد در صورت استفاده از رکوردهای مابین روزهای ۵ تا ۳۰۵ روزگی در آنالیزها تعداد ۳۰۱ PBV برای هر حیوان به دست می‌آید که این کار هم مستلزم صرف هزینه بوده و هم زمان‌بر است. بنابراین هدف تحقیق حاضر کاهش تعداد ارزش‌های اصلاحی برآورد شده با استفاده از برآورد مؤلفه اصلی می‌باشد. این تجزیه و تحلیل یعنی تعیین مؤلفه اصلی از روی پیش‌بینی ارزش‌های اصلاحی صفات تولیدی گاوهای

² Principal Component Analyze (PCA)

³ Principal Component (PC)

⁴ Predicted Breeding Value (PBV)

¹ Test Day Record (TDR)

شیری هلشتاین، برای اولین بار در ایران صورت می‌گیرد.

مواد و روش

داده‌های مورد استفاده

رکوردهای روز آزمون مربوط به صفات تولیدی گاوهای شیری هلشتاین خالص (گاوهای مورد مطالعه با توجه به شجره‌ای که داشتند ۱۰۰ درصد نژاد هلشتاین بودند) ایران از مرکز اصلاح نژاد کشور (واقع در کرج) تهیه گردید. تجزیه و تحلیل داده‌ها در آزمایشگاه بیوانفورماتیک و داده‌پردازی علوم دامی دانشکده کشاورزی دانشگاه تبریز انجام شد. تولید شیر مهم‌ترین صفت اقتصادی گاو شیری است. بنابراین برای تجزیه و تحلیل این رکوردها و برآورد پارامترهای ژنتیکی به مدلی با دقت بالا نیاز است. زیرا این تجزیه و تحلیل سهم قابل‌توجهی در افزایش بهره‌وری برنامه‌های انتخاب دارد (کوبوکی و همکاران ۲۰۱۱). صفات مورد مطالعه شامل: تولید شیر (۵ تا ۱۰ رکورد ماهانه به‌ازای هر گاو به‌طور متوسط ۸/۸ رکورد ماهانه)، درصد چربی شیر (۵ تا ۱۰ رکورد ماهانه و به‌طور متوسط ۸ رکورد ماهانه) و درصد پروتئین (۵ تا ۱۰ رکورد ماهانه و به‌طور متوسط ۷/۹ رکورد ماهانه) در بازه زمانی بین ۵ تا ۳۰۵ روز شیردهی، بود. مجموعه رکوردها از جمعیت گاوهای شیری هلشتاین ایران شکم زایش اول در دامنه سنی اولین زایش بین ۲۱ تا ۴۸ ماهگی برای سال‌های ۱۳۶۷ تا ۱۳۹۴ جمع‌آوری شد. به‌علاوه وجود حداقل یک رکورد ماهانه در ۹۰ روز اول بعد از زایش

برای گاو ضروری بود در غیر این صورت حذف می‌شود. سپس، رکوردهای تولید شیری که در دامنه ۱/۵ تا ۷۰ کیلوگرم، رکوردهای درصد چربی در دامنه بین ۱ تا ۶ درصد و رکوردهای درصد پروتئین در دامنه بین ۱ تا ۷ درصد، بودند در فایل داده باقی ماندند (محمدی و همکاران ۱۳۹۱). درنهایت برای صفت تولید شیر تعداد ۷۳۸۳۹، برای صفت درصد چربی تعداد ۶۵۱۶۵ و برای صفت درصد پروتئین تعداد ۴۶۸۸۱ گاو مورد استفاده قرارگرفت.

برآورد پارامترهای ژنتیکی

برای برآورد پارامترهای ژنتیکی از مدل رگرسیون تصادفی تک صفتی و برای محاسبه PBV از روش بیزی استفاده شد. مدل ماتریسی رگرسیون تصادفی به‌صورت زیر نوشته می‌شود:

$$y = Xb + Qa + Zpe + e$$

که در این مدل، y بردار مشاهدات، b بردار ضرایب اثرات ثابت که شامل سال تولد، گله روز آزمون و دفعات دوشش برای همه صفات در نظر گرفته شد، a بردار ضرایب رگرسیون برای اثرات تصادفی ژنتیک افزایشی، pe بردار ضرایب رگرسیون تصادفی برای اثرات تصادفی محیط دائم و e بردار اثرات باقی‌مانده را نشان می‌دهند. X ، Q و Z نیز ماتریس‌های طرح می‌باشند که به‌ترتیب ارتباط اثرات ثابت، ژنتیک افزایشی و محیط دائم را با بردار مشاهدات برقرار می‌کنند. مدل حیوانی روز آزمون رگرسیون تصادفی که اثر روز شیردهی در این مدل برای صفات موردبررسی تصحیح‌شده به‌قرار زیر است:

$$y_{tijkml} = HTD_i + BY_j + MT_k + \sum_{n=1}^3 AFC_{mn} \phi_n + \sum_{n=0}^2 a_{mn} \phi_n + \sum_{n=0}^4 pe_{mn} \phi_n + e_{tijkml}$$

ها هم از رویه CORR در محیط برنامه‌نویسی SAS استفاده شد. در نهایت با استفاده از همبستگی‌های برآورد شده و ویژه‌بردار و ویژه‌مقدار می‌توان PC را برای ارزش‌های اصلاحی صفات مورد مطالعه برآورد کرد.

تجزیه به مؤلفه اصلی

هدف اصلی PCA در این پژوهش، کاهش برآورد متغیرهای هر یک از صفات مورد مطالعه به تعدادی متغیر متعامد، به نام مؤلفه اصلی با از دست دادن حداقل اطلاعات می‌باشد. PCA از PBV استاندارد شده صفات مورد مطالعه (PBV تولید شیر، درصد چربی، درصد پروتئین) به دست آمد. فرمول استاندارد کردن PBV صفات مورد مطالعه به قرار زیر بود:

$$z = \frac{x - \bar{x}}{S}$$

که z ارزش استاندارد شده x ، $\bar{x}$ میانگین صفت مورد نظر و S انحراف معیار می‌باشد. برای استاندارد کردن PBV ها از نرم‌افزار SAS و رویه STANDARD استفاده شد.

مؤلفه اصلی توسط ترکیب خطی از متغیرهای اصلی استاندارد شده ویژه‌بردار محاسبه می‌شود. ارزش مطلق یک ویژه‌بردار اهمیت صفات را در PC ها تعیین می‌کند. هر ویژه‌بردار توسط مقدار ویژه‌ای از ماتریس همبستگی داده‌ها محاسبه می‌شود. از طرفی دیگر ویژه-مقدارها با واریانس هر PC مرتبط می‌باشد (رنجر ۲۰۰۲). هر ویژه‌مقدار با یک بردار واحد به نام ویژه-بردار مرتبط می‌باشد. ویژه‌بردار مقدار و جهت واریانس هر صفت را در مؤلفه‌های اصلی نشان می‌دهد. اولین مؤلفه اصلی (PC1) بیشترین درصد از کل واریانس ژنتیک افزایش را تبیین می‌کند. دومین مؤلفه اصلی (PC2) دومین بیشترین درصد از کل واریانس را تشریح می‌کند و به این ترتیب ادامه دارد تا زمانی که کل واریانس تبیین شود. بنابراین در یک مجموعه داده با p متغیر، مؤلفه اصلی i ام برابر است با:

$$PC_i = a_{i1}X_1 + a_{i2}X_2 + \dots + a_{ij}X_j$$

که در مدل فوق، y_{ijklm} ، t امین رکورد روز آزمون در زیرگروه‌های گله-سال-ماه رکورد روز آزمون^۱ (HTD) i ام، سال تولد (BY) j ام و دفعات دوشش (M) k ام مربوط به حیوان m ام؛ AFC_{mn} ، n امین ضریب رگرسیون ثابت برای سن اولین زایش حیوان m ام؛ a_{mn} و pe_{mn} ، به ترتیب n امین ضریب رگرسیون تصادفی اثر ژنتیک افزایشی و اثر محیط دائمی حیوان m ام؛ Φ_n ، چندجمله‌ای‌های لژاندر n ام برای روز t ام مرتبط با حیوان m ام؛ e_{ijklm} ، اثرات تصادفی باقی‌مانده می‌باشند. ساختار واریانس کوواریانس اثرات تصادفی مربوطه به صورت رابطه زیر می‌باشد.

$$\text{var} \begin{pmatrix} a \\ pe \\ e \end{pmatrix} = \begin{pmatrix} G \otimes A & 0 & 0 \\ 0 & I\sigma_p^2 & 0 \\ 0 & 0 & R \end{pmatrix}$$

که A ماتریس ضرایب خویشاوندی بین گاوها، G ماتریس کوواریانس ژنتیکی ضرایب رگرسیون تصادفی، $\otimes$ ضرب کرونگر، I ماتریس واحد، σ_p^2 واریانس اثرات محیط دائم و R ماتریس قطری واریانس باقی‌مانده می‌باشند. برای PBV از مدل رگرسیون تصادفی و روش بیزی از برنامه GIBSS3F90 (میسئل ۲۰۰۲) در محیط لینوکس استفاده شد تعداد دوره‌های نمونه‌گیری گیس ۳۰۰۰۰ انتخاب شد. در هر تجزیه و تحلیل ۵۰۰۰۰ دور اول به عنوان دوره‌های قلق‌گیری^۲ حذف شد و فواصل بین نمونه‌گیری‌ها ۱۰۰ در نظر گرفته شد. برای این کار از ضرایب چندجمله‌ای لژاندر (۳،۵) براساس یافته‌های محمدی (۱۳۹۱)، استفاده شد. بدین صورت که برای اثرات ژنتیک افزایش از چندجمله‌ای لژاندر درجه ۳ و برای اثرات محیط دائم از چندجمله‌ای لژاندر درجه ۵ استفاده شد. با استفاده از این ضرایب لژاندر و برنامه GIBSS3F90 خروجی به دست می‌آید که از این خروجی برای پیش‌بینی BV برای همه‌ی روزهای شیردهی یعنی از روز ۵ ام تا روز ۳۰۵ ام استفاده شد. در این مطالعه برای به دست آوردن PBV از نرم‌افزار SAS (۲۰۰۳) و رویه IML استفاده شد. همچنین برای به دست آوردن همبستگی بین PBV

¹ Herd Teat Day (HTD)

² Burn-in

$$SSC_{ij} = \frac{eigenvector_{ij}}{\sqrt{eigenvalue_j}}$$

که در آن SSC_{ij} ضریب نمره استاندارد شده صفت i ام و مؤلفه اصلی j ام برای PBV می‌باشد (بوزانسکاس و همکاران ۲۰۱۲). نمرات (شاخص) مؤلفه اصلی به شرح زیر محاسبه می‌شود:

$$PC_{jl} = \sum_{i=1}^l SSC_{ij} PBV_{il}$$

که در آن PC_{jl} نمره (شاخص) مؤلفه‌های اصلی برای حیوان l ام در مؤلفه اصلی j ام، SSC_{ij} ضریب نمره استاندارد شده برای صفت i ام و مؤلفه اصلی j ام برای PBV و PBV_{il} پیش‌بینی ارزش اصلاحی استاندارد شده i امین صفت برای l امین حیوان می‌باشد (بوزانسکاس و همکاران ۲۰۱۲).

($i=1,2,\dots,301$; $j=1,2,\dots,301$) که در آن a_{ij} j امین ویژه‌بردار و X_j هم j امین ارزش متغیر اصلی می‌باشد (ساواگناگو و همکاران ۲۰۱۱). این PC ژنتیکی توسط نرم‌افزار SAS و رویه PRINCOMP برآورد شده است (باتیستا پینتو و همکاران ۲۰۰۶).

با استفاده از PBV های استاندارد شده در این تجزیه-وتحلیل، هر مؤلفه اصلی می‌تواند یک مقدار جدیدی به نام نمره مؤلفه اصلی تولید نماید که از طریق مجموع ارزش اصلاحی استاندارد شده هر صفت توسط ضریب نمره استاندارد شده مربوطه ایجاد می‌شود. به این ترتیب، مؤلفه‌های اصلی می‌توانند به عنوان یک شاخص برای ارزیابی گاوها برای صفات مختلف استفاده شوند. ضریب نمره استاندارد هر PBV در هر مؤلفه اصلی با استفاده از فرمول زیر به دست می‌آید:

جدول ۱- آماره توصیفی صفات تولید شیر، درصد چربی و درصد پروتئین

Table 1- Descriptive statistics of milk yield, fat percent and protein percent traits

اطلاعات Information	تولید شیر Milk yield	درصد چربی Fat percent	درصد پروتئین Protein percent
تعداد حیوانات رکورددار No. of animals with record	73839	65165	46881
تعداد حیوانات موجود در شجره No. of animals in pedigree	176390	176390	176390
تعداد پدرها No. of sires	2429	2429	2429
تعداد مادرها No. of dams	86298	86298	86298
تعداد رکوردها No. of records	649089	523680	371175
تعداد گله No. of herd	365	385	254
تعداد گله-روز آزمون No. of HTD	3213	3087	2005
میانگین تولید (انحراف معیار) Means (S.D.)	32.11±8.72 (kg)	3.17±0.84 (%)	3.07±0.40 (%)

S.D. = Standard Deviation, HTD= Herd Test Day

جدول ۲- میانگین و انحراف معیار صفت تولید شیر، درصد چربی و درصد پروتئین در ماه‌های مختلف دوره شیردهی

Table 2- Means and standard deviations of milk yield, fat percent and protein percent traits at different months of lactation

روزهای مختلف شیردهی DIM	صفت Trait	درصد چربی (%) Fat percent (%)			تولید شیر (kg) Milk yield (kg)			درصد پروتئین (%) Protein percent (%)		
		انحراف معیار			انحراف معیار			انحراف معیار		
		تعداد No.	میانگین Mean	S.D.	تعداد No.	میانگین Mean	S.D.	تعداد No.	میانگین Mean	S.D.
5-35		52432	3.38	0.92	67681	29.65	0.92	36925	3.1	0.46
36-65		56530	3.8	0.84	70269	34.06	0.84	39672	2.93	0.38
66-95		55146	3.04	0.82	68302	34.61	0.82	39014	2.96	0.37
96-125		56965	3.05	0.82	70388	34.28	0.82	40351	3.005	0.36
126-155		54838	3.07	0.81	67743	33.65	0.81	39021	3.05	0.36
156-185		56186	3.12	0.81	68822	32.87	0.81	40031	3.08	0.37
186-215		53411	3.17	0.81	65110	31.93	0.81	38078	3.11	0.38
216-245		53264	3.24	0.81	65489	30.71	0.81	37994	3.14	0.39
246-275		47700	3.3	0.80	58972	29.31	0.80	33943	3.18	0.39
276-305		37208	3.37	0.82	46313	28.21	0.82	26146	3.22	0.40

S.D. = Standard Deviation, DIM= Days in Milk

نتایج

آماره توصیفی

مشخصات شجره‌ای و آماره توصیفی داده‌های پژوهشی حاضر در جدول ۱ ارائه شده است. همچنین در جدول ۲ میانگین و انحراف معیار صفات تولید شیر در ماه‌های مختلف دوره شیردهی گزارش شده است. با توجه به اطلاعات این دو جدول می‌توان چنین نتیجه گرفت که میانگین تولید شیر در ماه‌های مختلف شیردهی بین ۲۸/۲۱ تا ۳۴/۶۱ کیلوگرم، و برای درصد چربی بین ۳/۰۴ تا ۳/۳۸ درصد و برای درصد پروتئین نیز ۲/۹۳ تا ۳/۲۲ درصد بود. همچنین تنوع در رکوردهای صفت تولید شیر نسبت به دو صفت دیگر بیشتر بود.

همبستگی‌های ژنتیکی

با استفاده از ارزش‌های اصلاحی به‌دست‌آمده برای تمامی صفات، همبستگی ژنتیکی برآورد شد. برای این همبستگی ژنتیکی نمودار همبستگی ترسیم شد که در شکل‌های ۱ تا ۳ نشان داده شده است.

همبستگی ژنتیکی مربوط به قطر نمودارها که نشان‌دهنده همبستگی ژنتیکی هر BV با خود می‌باشد مثبت یک بوده و با رنگ خاکستری پررنگ در این اشکال قابل‌مشاهده است. از طرف دیگر همبستگی بین PBV های مربوط به اواسط دوره شیردهی در تمام صفات مثبت و بالا بود. همبستگی ژنتیکی درصد چربی هیچ‌کدام از PBV ها با یکدیگر منفی نبود؛ با این حال همبستگی بین PBV های روزهای مختلف هرچقدر که با PBV های اواسط دوره شیردهی فاصله بیشتری داشت، به صفر نزدیک‌تر بود. درحالی‌که برای دو صفت بعدی چنین نیست. با توجه به اشکال ۱ تا ۳ می‌توان دریافت که همبستگی بین اواسط دوره شیردهی در هر سه صفت مورد مطالعه بیشتر می‌باشد و دامنه تغییرات همبستگی ژنتیکی صفات تولید شیر و درصد پروتئین بین ۱ تا ۱- بود، از طرفی هم هرچه همبستگی بین PBV های روزهای مختلف بالا باشد، دقت هم زیاد می‌شود. بنابراین دقت PBV های مربوط به اواسط دوره

شیردهی بیشتر از اوایل و اواخر دوره شیردهی می‌باشد.

تجزیه و تحلیل مؤلفه‌های اصلی برای ارزش اصلاحی صفات مورد مطالعه

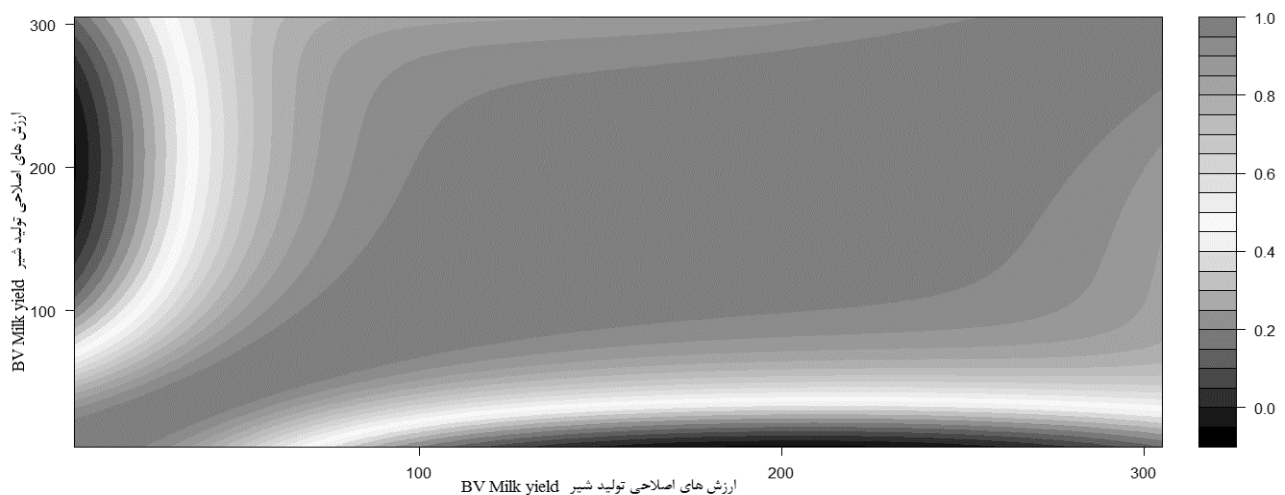
تجزیه و تحلیل مؤلفه اصلی برای PBV صفات تولید شیر، درصد چربی و پروتئین شیر با استفاده از PBV روزانه استاندارد شده هر صفت برآورد شد و PBV هایی تولید شیر که با اولین PC بیشترین ارتباط را داشتند در جدول ۳ آمده است. همچنین نمودار PC برای PBV درصد چربی و PBV درصد پروتئین در شکل-های ۴ و ۵ آمده است. برای صفت تولید شیر PC1، ۹۹/۴۸ درصد از کل واریانس ژنتیکی را تبیین کرد. همین‌طور برای صفت درصد چربی شیر PC1، ۸۸/۶۹ درصد و PC2، ۹/۸ درصد و برای صفت درصد پروتئین شیر هم PC1، ۶۶/۸ درصد و PC2، ۳۳/۲ درصد از کل واریانس ژنتیکی را تبیین کردند.

جدول ۳ نشان‌دهنده این است که PC1 مربوط به صفت تولید شیر با همه ارزش‌های اصلاحی به‌جز روز ۵۶ ام و روز ۲۳۱ ام، همبستگی ژنتیکی بالای ۰/۰۵۶ (بزرگ‌تر از ۰/۰۵۶ یا کوچک‌تر از ۰/۰۵۶-) دارد. همچنین با توجه به شکل‌های ۴ و ۵ می‌توان همبستگی ژنتیکی بالای بین هرکدام از صفات را با اولین و دومین مؤلفه اصلی مشاهده کرد. بنابراین، می‌توان چنین برآورد کرد که اولین و دومین مؤلفه اصلی صفت درصد چربی با ارزش‌های اصلاحی روزهای ۸۳ ام تا ۲۲۲ ام و ۵ ام تا ۳۱ ام به‌ترتیب با اثرات ۰/۰۶ (بزرگ‌تر از ۰/۰۶ یا کوچک‌تر از ۰/۰۶-) و ۰/۱ (بزرگ‌تر از ۰/۱ یا کوچک‌تر از ۰/۱-) بیشترین ارتباط را دارا می‌باشند؛ اما بیشترین همبستگی ژنتیکی بین صفت درصد پروتئین با اولین مؤلفه اصلی با ارزش‌های اصلاحی روز ۹۹ ام تا روز ۱۶۸ و روز ۲۸۹ ام تا روز ۳۰۵ با ضریبی بالای ۰/۰۷ (بزرگ‌تر از ۰/۰۷ یا کوچک‌تر از ۰/۰۷-) و دومین مؤلفه اصلی با ارزش‌های اصلاحی روزهای ۴۰ ام تا ۶۲ ام و ۲۲۱ ام تا ۲۳۵ ام بیشترین همبستگی ژنتیکی را با ضریب بالای ۰/۰۹ (بزرگ‌تر از ۰/۰۹ یا کوچک‌تر از ۰/۰۹-) دارد.

نمره استاندارد شده به‌دست آمد و با ضرب این ضریب به ارزش‌های اصلاحی روزانه هر گاو، نمره مؤلفه اصلی برای همان حیوان برآورد شد. درنهایت برای ارزیابی گاوهای شیری می‌توان از فرمول نمره مؤلفه اصلی همانند فرمول زیر استفاده کرد. برای مثال، فرمول نمره مؤلفه اصلی ارزش اصلاحی تولید شیر برای هر گاو بدین قرار است:

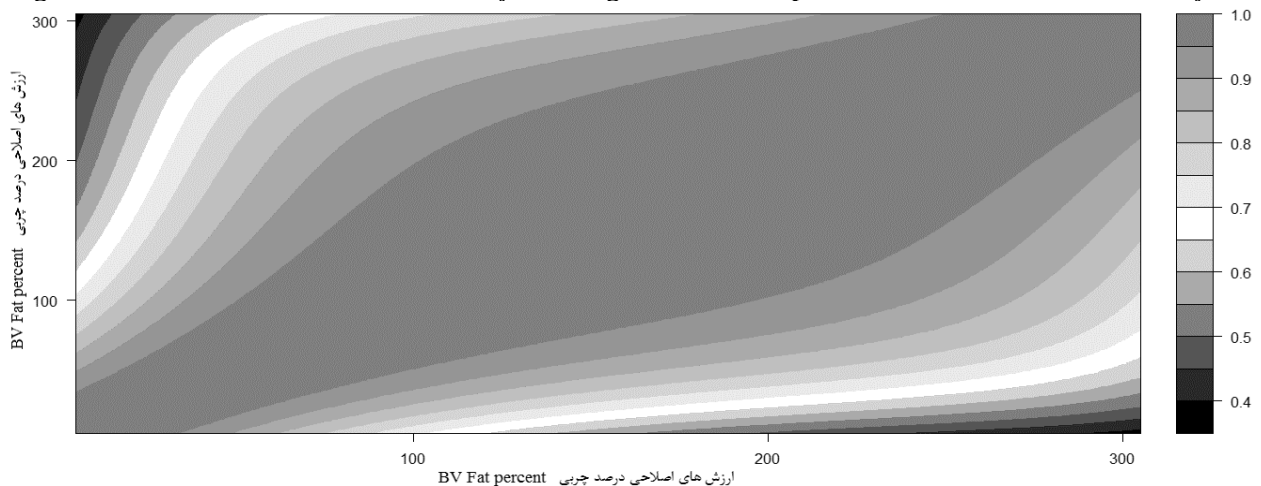
$$PC_{1j} = 0.00335BV_{98l} + 0.00335BV_{99l} + \dots - 0.00335BV_{304l} - 0.00335BV_{305l}$$

گاوهای شیری هلشتاین را می‌توان با استفاده از شاخص انتخابی که توسط نمره مؤلفه اصلی برای هرکدام از صفات انتخاب نمود. برای انجام این کار با استفاده از فرمول‌های توضیح داده‌شده و با استفاده از نرم‌افزار SAS و رویه IML، ویژه‌مقدار روزانه هر مؤلفه اصلی برای هر روز، برریشه دوم ویژه‌بردار همان مؤلفه اصلی تقسیم شد و بدین‌صورت ضریب



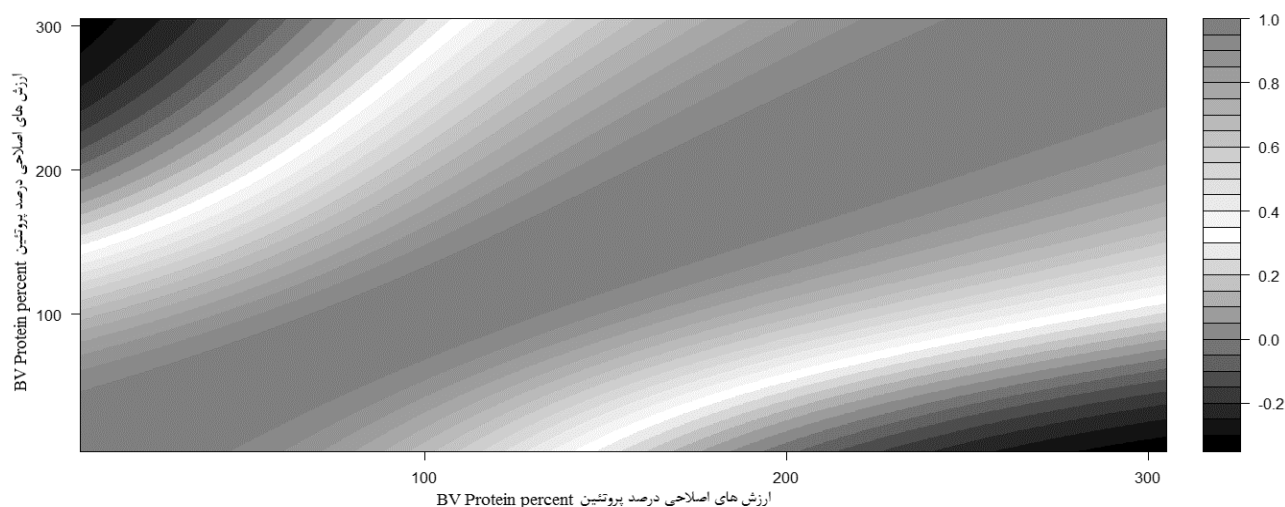
شکل ۱- همبستگی ژنتیکی بین ارزش‌های اصلاحی روزانه تولید شیر گاو شیری هلشتاین ایران

Figure 1. Genetic correlations between daily breeding values of milk production of Iranian Holstein dairy cattle



شکل ۲- همبستگی ژنتیکی بین ارزش‌های اصلاحی روزانه درصد چربی گاو شیری هلشتاین ایران

Figure 2- Genetic correlations between daily breeding values of fat percent trait of Iranian Holstein dairy cattle



شکل ۳- همبستگی ژنتیکی بین ارزش‌های اصلاحی روزانه درصد پروتئین گاو شیری هلشتاین ایران

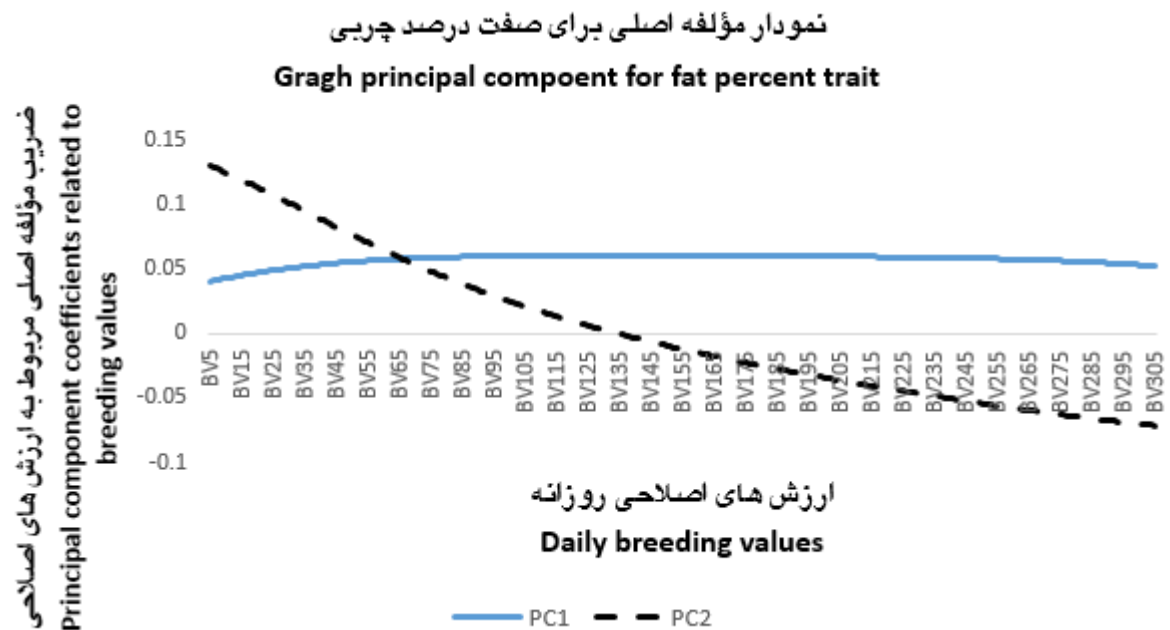
Figure 3- Genetic correlations between daily breeding values of protein percent of Iranian Holstein dairy cattle

جدول ۳- ضریب مؤلفه اصلی (ویژه مقدار) مربوط به ارزش‌های اصلاحی استاندارد شده صفت تولید شیر

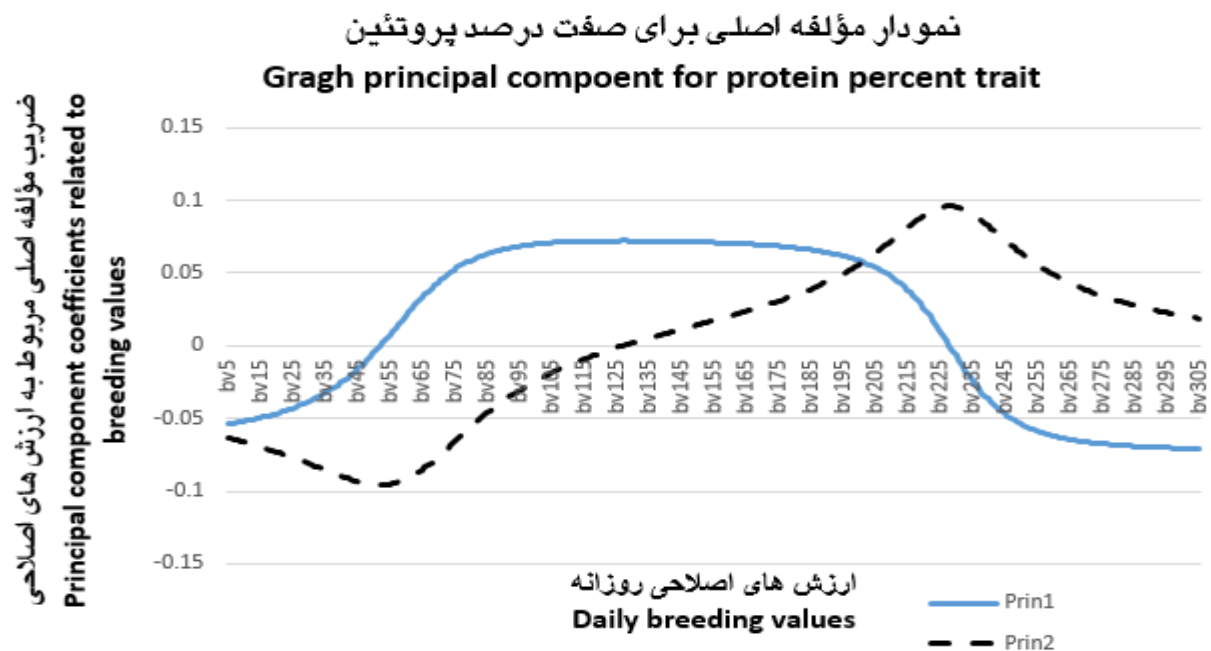
Table 3- Principal component coefficients (eigenvectors) related to standardized breeding values of the milk yield trait

ارزش‌های اصلاحی BV _s	اولین مؤلفه اصلی PC1	ارزش‌های اصلاحی BV _s	اولین مؤلفه اصلی PC1
BV98 – BV191	0.05779	.	.
BV82 – BV97 & BV192 – BV208	0.057789	BV53	-0.05765
BV77 – BV81 & BV209 – BV213	0.057788	BV234	-0.05770
BV74 – BV76 & BV214 – BV216	0.057787	BV52	-0.05771
BV72, BV73 & BV217	0.057786	BV51 & BV235	-0.05774
BV71, BV218 & BV219	0.057785	BV50	-0.05775
BV70	0.057784	BV49, BV236	-0.05776
BV69 & BV220	0.057783	BV47 & BV48	-0.05777
BV68 & BV221	0.057782	BV237 & BV238	-0.05778
BV67 & BV222	0.057781	BV39 – BV46 & BV239 – BV242	-0.05779
.	.	BV5 – BV38 & BV244 – BV305	-0.05779
.	.		

PC= Principal Component, BV= Breeding value



شکل ۴- نمودار خطی ضریب مؤلفه اصلی مربوط به ارزش‌های اصلاحی استاندارد شده صفت درصد چربی
Figure 7- Principal component coefficients related to standardized breeding values of the fat percent trait



شکل ۵- نمودار خطی ضریب مؤلفه اصلی مربوط به ارزش‌های اصلاحی استاندارد شده صفت درصد پروتئین
Figure 8- Principal component coefficients related to standardized breeding values of the protein percent trait

بحث

در بین مدل‌هایی که براساس روز آزمون بررسی می‌شوند، مدل رگرسیون تصادفی جهت افزایش صحت تخمین EBV^۱، نسبت به مدل‌های دیگر، به دلیل داشتن مزیت‌هایی همچون امکان منظور نمودن اثر محیطی خاص برای هرروز رکوردگیری در برآورد BV دام‌ها، افزایش دقت برآورد پارامترهای ژنتیکی، اخیراً در ارزیابی‌های ژنتیکی گاوهای شیری در بیشتر کشورها مورد استفاده قرار می‌گیرد (محمدی و همکاران ۱۳۹۲ و سترابل و همکاران ۲۰۰۴). اولین کاربرد مدل رگرسیون تصادفی در ارزیابی ژنتیکی گاوهای شیری، استفاده از رکوردهای روز آزمون انفرادی برای برازش منحنی شیردهی بود (بهلولی و همکاران ۱۳۹۱، شفر و دکرز ۱۹۹۴ و جامروزیک و همکاران ۱۹۹۷). با این که مدل رگرسیون تصادفی یکی از محبوب‌ترین مدل‌ها برای تجزیه و تحلیل ژنتیکی صفات می‌باشد، به دلیل حجم محاسباتی بالا محققان درصدد یافتن راهی برای کاهش این حجم محاسباتی می‌باشند. بنابراین یکی از شایع‌ترین این روش‌ها استفاده از مؤلفه‌های اصلی می‌باشد. در این زمینه تحقیقاتی مختلفی انجام شده است.

ایگیدلو-گومز و همکاران (۲۰۱۵) با استفاده از یک مدل تک صفتی، صفات تولید شیر در ۲۷۰ روز، سن اولین زایش، افزایش وزن قبل از شیرگیری و وزن در سنین یک‌سالگی، ۱۸ ماهگی و ۲۴ ماهگی را بررسی کردند. این محققین تجزیه و تحلیل مؤلفه اصلی را با استفاده از ارزش‌های اصلاحی صفات فوق بر روی اصلاح نژاد ۱۳۳ گاومیش نر منتخب از ۹۶۱ گاومیش نر را انجام دادند. در تحقیق فوق معلوم شد که سه PC نخست ۶۵/۷۸٪ از واریانس کل برای EBV صفات فوق را تشکیل می‌دهند. اولین PC با صفات افزایش وزن قبل از شیرگیری، وزن یک‌سالگی، وزن ۱۸ ماهگی و وزن دوسالگی رابطه مثبت بالایی داشت. دومین PC با اثرات ژنتیک مادری بر روی صفات افزایش وزن قبل از شیرگیری، وزن یک سالی و صفت تولید شیر در ۲۷۰ روز رابطه مثبتی داشت. سومین PC با صفات تولید

شیر در ۲۷۰ روزگی و سن اولین زایش رابطه مثبت و بالایی داشت.

بولیگون و همکاران (۲۰۱۳) نیز درمجموع از ۶۱۵۲۸ رکورد وزنی از ۲۲۲۴۶ رأس گاو نلور^۲ که بین سال‌های ۱۹۸۴ تا ۲۰۰۲ به دنیا آمده بودند برای مقایسه روش‌های تجزیه و تحلیل چند صفت مختلف برای وزن تولد تا وزن بلوغ مورد ارزیابی قرار گرفتند. مدل‌هایی که در این مطالعه مورد استفاده قرار گرفتند شامل: مدل چند صفتی استاندارد، مدل‌های PC ها و مدل‌های فاکتور تجزیه و تحلیل می‌باشد. با توجه به معیارهای انتخاب مدل، مدل سه PC بدون نیاز به مدل‌های فاکتور تجزیه و تحلیل، بهترین و مناسب‌ترین مدل بود. برآورد (کو) واریانس فنوتیپی، ژنتیک افزایشی مستقیم و مادری، محیط دائم مادری و باقی‌مانده با استفاده از مدل چند صفتی استاندارد و مدل سه PC به دست آمد.

بوزانسکاس و همکاران (۲۰۱۲) از اطلاعات فنوتیپی گاوهای نر نژاد کنکیم^۳ برای به دست آوردن پارامترهای ژنتیکی در صفات تولیدمثلی و رشد با استفاده از یک مدل مختلط خطی دام استفاده کردند. علاوه بر این برای این صفات با استفاده از PCA روابط بین BV های برآورد شده حیوان (EBVs) را کشف کردند. صفات مورد مطالعه این تحقیق شامل: سن اولین زایش، سن در زایش دوم، فاصله زایش و وزن بدن در سن ۴۲۰ روزگی بود. در این مطالعه دو PC نخست ۷۳/۳۷٪ از واریانس کل را تشکیل دادند. اولین مؤلفه با صفات سن در اولین و دومین زایش و فاصله زایش رابطه مثبت و بالایی داشت ولی دومین مؤلفه فقط با صفت افزایش وزن رابطه مثبت و بالایی داشت.

باتیستا پیتو و همکاران (۲۰۰۶) روش مؤلفه‌های اصلی را برای تجزیه و تحلیل صفات عملکردی و لاشه مرغ مورد استفاده قرار دادند؛ که برای این کار از جمعیت ۳۷۴۲ نطفه مرغ اهلی نمونه‌گیری کردند. صفات عملکردی شامل وزن بدن از سن ۳۵ تا ۴۲ روزگی، افزایش وزن در سن ۳۵ تا ۴۲ روزگی بود و صفات

^۲ Nellore^۳ canchim^۱ Estimated Breeding Value (EBV)

آن‌ها برای هدف ارزیابی و مقایسه حیوانات استفاده می‌شود.

این مطالعه برای اولین بار در ایران به دنبال راهی برای کاهش PBV های محاسبه‌شده برای صفات تولیدی گاو هلشتاین ایران است. به خاطر همین از مؤلفه‌های اصلی استفاده شد. با توجه به نتایج به‌دست‌آمده از تجزیه و تحلیل مؤلفه‌های اصلی ارزش‌های اصلاحی استاندارد شده صفات تولید شیر، مشاهده شد که اولین مؤلفه‌های اصلی درصد بالایی از واریانس کل را تشکیل می‌دهند.

نتیجه‌گیری کلی

با توجه به اینکه هزینه رکوردبرداری بالا بوده و از طرفی با در نظر گرفتن همبستگی بالا بین پیش‌بینی ارزش‌های اصلاحی به‌نظر می‌رسد، نیازی نیست برای همه روزها ارزش‌های اصلاحی را پیش‌بینی نمود و به طبع آن بتوان رکوردبرداری را کاهش داد. همچنین با توجه به ارتباط بین مؤلفه‌های اصلی با برخی از روزهای ارزش‌های اصلی پیشنهاد می‌شود که از مؤلفه‌های اصلی برای هرکدام از صفات مورد مطالعه برای شاخص انتخاب بهره برد.

سپاسگزاری

از مرکز اصلاح نژاد کرج و تمام افرادی که در نگارش این مقاله کمک کردند، تشکر و قدردانی می‌گردد.

لاشه هم شامل: وزن کبد، قلب، سنگدان، بال‌ها، ران‌ها، سینه و ریه بود. در این مطالعه پنج PC نخست حدوداً ۹۳٪/۳ از واریانس کل را تشکیل دادند. همچنین این محققین به این نتیجه رسیدند که آنالیز مؤلفه‌های اصلی امکان درک بهتری از پیچیدگی همبستگی بین صفات و کاهش تعدادی از صفات را فراهم می‌سازد، تنها با استفاده از چند PC می‌توان با از دست دادن اطلاعات اندک به نتیجه مطلوبی رسید.

باتیستا پینتو و همکاران (۲۰۰۶) گزارش کردند ضریبی که با استفاده از دو مؤلفه اصلی اول برای شاخص انتخاب برآورد شد به ۱/۷۵ رسید این درحالی‌که بود که ضریب شاخص انتخاب با استفاده از همه صفات، ۱/۷۷ بود. این نتایج نشان داد که ممکن است بتوان افزایش وزن ژنتیکی به‌دست‌آمده از شاخص انتخابی که از PC ها استفاده شده است را با شاخص انتخابی که در آن از تمام صفات استفاده‌شده، مقایسه نمود. با این حال، علی‌رغم مزیت‌های بالقوه استفاده از مؤلفه‌های اصلی در ارزیابی‌های چند متغیره هنوز از مؤلفه‌های اصلی در اصلاح نژاد بسیار کم استفاده شده است. شاید دلیل آن تا حدودی به این مسئله برمی‌گردد که چند مؤلفه اصلی اولیه همیشه نمی‌توانند بخش زیادی از واریانس کل را تشریح کند. با این وجود، در شرایط مناسب استفاده عملی آن می‌تواند تا حدود زیادی ارزیابی حیوانات را در برنامه‌های اصلاح نژادی تسهیل دهد، مزایای زیادی به‌دست می‌آید (باتیستا پینتو و همکارانش ۲۰۰۶). با این وجود با توجه به نتایج به‌دست‌آمده در این پژوهش دو مؤلفه اصلی اولیه درصد بالایی از واریانس کل را برای هرکدام از صفات تبیین می‌نماید بنابراین می‌توان از مؤلفه اصلی برای افزایش بهره ژنتیکی بهره جست.

با توجه به مقاله ایب^۱ (۱۹۸۹) به‌دست آوردن شاخص انتخاب به‌صورت دیگر امکان‌پذیر شده که جالب بودند. در این مورد شاخص انتخاب از PC ها استفاده‌شده بود که می‌توانست به‌تنهایی توسط پنج ضریب وزنی به‌دست آید که در مقایسه با شاخصی که شامل همه صفات است کاری بسیار آسانی می‌باشد. از این رو از

^۱ Ibe

منابع مورد استفاده

- Agudelo-Gómez D, Pineda-Sierra S and Cerón-Muñoz MF, 2015. Genetic evaluation of Dual-Purpose buffaloes (*Bubalus bubalis*) in Colombia using principal component analysis. *PloS one* 10(7): e0132811.
- Aziz MA, Nishida S, Suzuki K and Nishida A, 2005. Estimation of direct and maternal genetic and permanent environmental effects for weights from birth to 356 days of age in a herd of Japanese Black cattle using random regression. *Animal Science* 83(3): 519-530.
- Batista Pinto LF, Packer IU, de Melo CMR, Ledur MC and Coutinho LL, 2006. Principal components analysis applied to performance and carcass traits in the chicken. *Animal Research* 55(5): 419-425.
- Bignardi AB, El Faro L, Cardoso VL, Machado PF and de Albuquerque LG, 2009. Random regression models to estimate test-day milk yield genetic parameters Holstein cows in Southeastern Brazil. *Livestock Science* 123(1): 1-7.
- Bohlouli M, Alijani S, Shodja J, Eghbal A and Pirani N, 2012. Genetic parameters and genotype by environment interaction in Iranian Holstein dairy cattle. *Animal science researches* 22(3): 149-159.
- Bohlouli M, Shodja J, Alijani S, Eghbal A, 2013. The relationship between temperature-humidity index and test-day milk yield of Iranian Holstein dairy cattle using random regression model. *Livestock Science* 157(2): 414-420.
- Boligon AA, Bignardi AB, Mercadante MEZ, Lobo RB and Albuquerque LG, 2013. Principal components and factor analytic models for birth to mature weights in Nellore cattle. *Livestock Science* 152(2): 135-142.
- Buzanskas ME, Savegnago RP, Grossi DA, Venturing GC, Queiroz SA, Silva LOC, Torres Junior RAA, Munari DP and Alencar MM, 2013. Genetic parameter estimates and principal component analysis of breeding values of reproduction and growth traits in female Canchim cattle. *Reproduction, Fertility and Development* 25(5): 775-781.
- Cobuci, JA, Costa CN, Neto JB and deFreitas AF, 2011. Genetic parameters for milk production by using random regression models with different alternatives of fixed regression modeling. *Revista Brasileira de Zootecnia* 40(3): 557-567.
- Ibe SN, 1989. Measures of size and conformation in commercial broilers. *Animal Breeding and Genetic* 106(1-6): 461 – 469.
- Jamrozik J, Schaeffer LR and Dekkers JCM, 1997. Genetic evaluation of dairy cattle using test day yields and random regression model. *Dairy Science* 80(6): 1217-1226.
- Jensen J, 2001. Genetic evaluation of dairy cattle using test-day models1. *Dairy Science* 84(12): 2803-2812.
- Kheirabadi K, Alijani S, Rafat SA and Moghaddam G, 2013. Estimation of genetic parameters of production traits and persistency of Iranian Holstein cows by random regression. *Management system* 1(3): 51-63.
- Kranis ASG, Sorensen D and Wolliams JA, 2007. The application of random regression models in the genetic analysis of monthly egg production in turkeys and a comparison with alternative longitudinal models. *Poultry Sciences* 86(3):470-475.
- Lewis RM and Brotherstone S, 2002. A genetic evaluation of growth in sheep using random regression techniques. *Animal Science* 74(1): 63-70.
- Misztal I, 2002. Gibbs3f90. <http://nce.ads.uga.edu/~ignacy/newprograms.html>.
- Mohammadi A, Alijani S, Rafat SA and Taghizadeh A, 2013. Comparison of fitting performance of random regression animal and sire models for yield traits of Iranian Holstein dairy cattle. *Animal science researches* 23(4): 159-178.
- Mohammadi A, Alijani S, Rafat SA, Taghizadeh A and Bohlouli M, 2012. Comparison of Fitting Performance of Polynomial Functions in Random Regression Model for Test Day Milk Yield in of Iranian Holstein Dairy Cattle. *Research on animal production (Scientific and Research)* 3(6): 47-63.
- Rencher AC, 2002. *Methods of Multivariate Analysis*. Wiley-Interscience, New York, NY.
- SAS Institute Inc, 2003. *SAS 9.1.3 Help and Documentation*, Cary, NC: SAS Institute Inc.
- Savegnago RP, Caetano SL, Ramos SB, Nascimento GB, Schmidt GS, Ledur MC and Munari DP, 2011.

- Estimates of genetic parameters, and cluster and principal components analyses of breeding values related to egg production traits in a White Leghorn population. *Poultry Science* 90(10): 2174 – 2188.
- Schaeffer LR and Dekkers JCM, 1994. Random regressions in animal models for test-day production in dairy cattle. *Proc. 5th World Congress of Genetics Applied to Livestock Production*. Guelph Ontario Canada XVIII: 443-446.
- Schaeffer LR, 2004. Application of random regression models in animal breeding. *Livestock Production Science* 86(1): 35-45.
- Strabel T, Ptak E and Szyda J, 2004. Multiple-lactation random regression test-day model for Polish Black and White cattle. *Interbull Bulletin* 32: 133-136.
- Takma C and Akbas Y, 2009. Comparison of fitting performance of random regression models to test day milk yields in Holstein Friesians. *Kafkas University Veterinary Fakultesi Dergisi* 15(2): 261-266.

The implementation of principal component analysis in genetic evaluation of Iranian dairy cattle

A Esmaili Nojadeh¹, S Alijani^{2*}, K Hasanpor³ and A Javanmard³

¹MSc Student, Department of Animal Science, Faculty of Agriculture, University of Tabriz, Tabriz, Iran

²Associate Professor, Department of Animal Science, Faculty of Agriculture, University of Tabriz, Tabriz, Iran

³Assistant Professor, Department of Animal Science, Faculty of Agriculture, University of Tabriz, Tabriz, Iran

*Coressponding author: Email: sad-ali@tabrizu.ac.ir

Introduction: In the quantitative genetics area, random regression model is one of the most accurate models for estimating daily breeding value in dairy cattle. However, because of the higher number records per each cow, application of this model is labor and time consuming. In addition, breeding values of cows at different days of lactation are highly correlated. The main objectives of the current study were to determine the relative importance of each breeding value at different days of lactation and to estimate the genetic principal components for the breeding values of Iranian Holstein dairy cattle for milk production traits.

Maternal and methods: records of milk production traits of first-parity dairy cows. Milk yield, fat percentage and protein percentage test-day records of 73839, 65165 and 46881 cows, respectively, from 230 herds with 176390 cows in their pedigree were used in the analyses. Only test-day records belonging to 5 to 305 days of lactation were used. The data belonged to cows were born between 1988 and 2015 with age at first calving ranged between 21 to 48 m. In addition, the existence of at least one monthly record in the first 90 days after calving was essential for the cow, otherwise it would be eliminated. These data were collected by National breeding center, Karaj, Iran. Genetic parameters were estimated by a random regression model and Bayesian approach using GIBSS3F90 software. The estimated breeding values at all days of lactation were calculated and standardized using the standard score (z). Then, Correlation matrices among breeding values at different days of lactation and genetic principal components of breeding values were estimated by PROC CORR and PROC PRINCOMP of SAS software, respectively. Finally, we could calculate principal component score as a selection criterion (selection index) for the selection of dairy cattle. For this purpose, the standardized score coefficient was obtained by dividing the daily eigenvector of each principal component by square root of its eigenvalue. The principal component score were calculated of the sum of the multiply between standardized score coefficient and daily standardized breeding values for each cow. However, the principal components could be used as an index to multiple traits evaluation of animals.

Results and discussion: The genetic correlations matrix between the estimated breeding values at different days of lactation demonstrated that the breeding values at the middle stage of lactation were highly correlated with the breeding values at the reaming stages of lactation. The genetic principal component analysis revealed that the first two principal components accounted for a high percent of total genetic variance of all studied traits. For milk yield, the first principal component explained 99.48% of genetic variance, while two first components explained almost 98.19% and 100% of genetic variance for fat percent and protein percent traits, respectively. The absolute value of correlations between the first principal component of milk yield and all breeding values (except for day 56 and day 231) were more than 0.056. The absolute values of correlations between the first principal component of fat percent and the daily breeding values were greater than 0.06 for days between 83 and 222; and for protein percent were greater than 0.07 for days 99 to 168 and days 289 to 305.

Considering the high correlation between breeding values seem to, were estimated breeding values for all days is not required. The first principal component milk yield trait with nearly all estimated breeding values, high correlation and first two principal component fat percent trait of estimated breeding values in the early and middle of lactation period had a high relationship. But first two principal component protein percent trait of estimated breeding values in the middle and later of lactation period had a high correlation.

Conclusions: Considering the high cost of recording system in dairy cattle industry and the high correlation between the breeding values, it seems that there is no need to predict the breeding value for all days of lactation. In other words, reducing the number of records per each cow may be beneficial at both economic and genetics stand points. Furthermore, due to the high, direct correlation between the principal components and daily breeding values, the implementation of principal components in the genetic merit evaluation of selection candidates for production-related traits is suggested.

Keywords: Genetic correlation, Principal component, Random regression