

Валидация первого метода прогнозирования молочной продуктивности первотелок по комплексному баллу линейных признаков

В. В. Гарт[✉], С. Г. Куликова, Е. В. Камалдинов, А. Ф. Петров, К. Н. Нарожных
Новосибирский государственный аграрный университет, Новосибирск, Россия
[✉]E-mail: gvlvl@yandex.ru

Аннотация. Цель исследования – валидация метода ранней оценки молочной продуктивности первотелок голштинской породы, основанного на использовании комплексного балла, рассчитанного методом последовательного суммирования скорректированных оценок линейных признаков по убыванию их индивидуальной корреляции с общим показателем продуктивности. **Методы.** Оценка качества модели производилась методом кросс-валидации на основе первичных данных, включающих комплексный балл, удой и выход жира за 305 дней. Оценка предсказательной способности комплексного балла производилась на уровне отдельных животных (путем расчета корреляции) и на уровне групп с одинаковым комплексным баллом экстерьера (с применением кросс-валидации полиномиальных моделей, построенных на средних значениях по группам). Такой подход позволил всесторонне оценить возможности использования комплексного балла экстерьера для непрямой оценки молочной продуктивности в практической селекционной работе. **Результаты исследования.** Установлено, что полиномиальные модели, построенные для средних значений продуктивности групп животных, агрегированных на основании одинакового комплексного балла, имеют высокую прогностическую способность. Так, кросс-валидированный R^2 для удоя равнялся 0,7566 при статистических ошибках RMSE 257,87 кг и MAE 177,84 кг, а для выхода жира кросс-валидированный R^2 составил 0,9688 при RMSE и MAE 3,61 кг и 2,78 кг соответственно. Результаты Post-hoc-анализа показали, что расчетная статистическая мощность полиномиальных моделей третьей степени для средних значений продуктивности была очень высокой (100 % для удоя и выхода жира) и статистически достоверной ($\alpha < 0,001$). Метод непрямой оценки и прогнозирования среднего уровня молочной продуктивности групп первотелок, объединенных по значению комплексного балла, может иметь практическое применение при оценке потенциала партий ремонтного молодняка, а также для принятия общих управленческих решений на уровне стада. **Научная новизна** проведенных исследований заключается в валидации нового способа раннего прогнозирования продуктивности первотелок голштинской породы по комплексу экстерьерных показателей.

Ключевые слова: голштинская порода, первотелки, прогнозирование продуктивности, удой, молочный жир, Западная Сибирь, экстерьер, линейная оценка, комплекс линейных признаков, кросс-валидация, регрессионная модель

Благодарности. Исследование выполнено по теме Государственного задания № FESF-2023-0016.

Для цитирования: Гарт В. В., Куликова С. Г., Камалдинов Е. В., Петров А. Ф., Нарожных К. Н. Валидация первого метода прогнозирования молочной продуктивности первотелок по комплексному баллу линейных признаков // Аграрный вестник Урала. 2025. Т. 25, № 09. С. 1372–1383. <https://doi.org/10.32417/1997-4868-2025-25-09-1372-1383>. (In Russ.)

Дата поступления статьи: 10.07.2025, **дата рецензирования:** 28.07.2025, **дата принятия:** 12.08.2025.

Validation of the first method for predicting milk productivity in first-calf heifers based on a composite score of linear type traits

V. V. Gart[✉], S. G. Kulikova, E. V. Kamaldinov, A. F. Petrov, K. N. Narozhnykh

Novosibirsk State Agrarian University, Novosibirsk, Russia

[✉]E-mail: gylvl@yandex.ru

Abstract. The purpose of this study was to validate a method for the early assessment of milk productivity in Holstein first-calf heifers. This method is based on a composite score calculated by sequentially summing the adjusted scores of linear type traits, ranked in descending order of their individual correlation with a total productivity indicator. **Methods.** The model quality was assessed using cross-validation on primary data, which included the composite score, 305-day milk yield, and fat yield. The predictive ability of the composite score was evaluated at two levels: for individual animals (by calculating correlation) and for groups of animals with the same exterior composite score (by applying cross-validation to polynomial models based on group mean values). This comprehensive approach allowed for a thorough assessment of the potential of using the exterior composite score for the indirect evaluation of milk productivity in practical breeding work. **Results.** It was established that polynomial models developed for the mean productivity values of animal groups, aggregated by their composite score, demonstrate high predictive ability. Specifically, the cross-validated R^2 for milk yield was 0.7566, with statistical errors RMSE of 257.87 kg and MAE of 177.84 kg. For fat yield, the cross-validated R^2 reached 0.9688, with RMSE and MAE of 3.61 kg and 2.78 kg, respectively. Post-hoc power analysis revealed that the statistical power of the third-degree polynomial models for mean productivity values was very high (100 % for both milk and fat yield) and statistically significant ($p < 0.001$). The method of indirect assessment and prediction of the average milk productivity level for groups of first-calf heifers, unified by their composite score, can be practically applied to evaluate the potential of replacement heifer batches and to make general management decisions at the herd level. **The scientific novelty** of this research lies in the validation of a new method for the early prediction of productivity in Holstein first-calf heifers based on a complex of exterior indicators.

Keywords: Holstein breed, first-calf heifers, productivity prediction, milk yield, milk fat yield, Western Siberia, conformation, linear type traits, composite score, cross-validation, regression model

Acknowledgements. The study was carried out on the topic of the State Assignment No. FESF-2023-0016.

For citation: Gart V. V., Kulikova S. G., Kamaldinov E. V., Petrov A. F., Narozhnykh K. N. Validation of the first method for predicting milk productivity in first-calf heifers based on a composite score of linear type traits. *Agrarian Bulletin of the Urals*. 2025; 25 (09): 1372–1383. <https://doi.org/10.32417/1997-4868-2025-25-09-1372-1383>. (In Russ.)

Date of paper submission: 10.07.2025, **date of review:** 28.07.2025, **date of acceptance:** 12.08.2025.

Постановка проблемы (Introduction)

Эффективная селекция является ключевым фактором повышения рентабельности молочного животноводства и обеспечения продовольственной безопасности. Основным критерием селекции традиционно выступает молочная продуктивность, оцениваемая по удою и качественным показателям молока (содержание жира, белка) за определенный период лактации (обычно 305 дней) или за несколько полных лактаций [1].

Однако прямая оценка продуктивности требует завершения значительной части первой лактации, что откладывает принятие решений по отбору и выбраковке животных, увеличивает затраты на их содержание и замедляет темпы селекционного про-

гресса. В связи с этим активно ведутся поиски методов непрямой оценки продуктивности, позволяющих прогнозировать будущие показатели в более ранние сроки выращивания или хозяйственного использования животных [2; 3].

Экстерьер животного отражает его конституциональные особенности, функциональное состояние и потенциал продуктивного долголетия. В частности, линейная оценка позволяет стандартизованно описывать признаки экстерьера, связанные с развитием молочной железы, конечностей, туловища, которые опосредованно могут влиять на способность животного к высокой продуктивности и продолжительности хозяйственного использования. Одним из перспективных направлений яв-

ляется использование признаков телосложения в качестве прогностических показателей молочной продуктивности [4–7].

Многочисленные исследования показали, что корреляция между отдельными признаками линейной оценки экстерьера и молочной продуктивностью, как правило, является низкой [8–11]. Это ограничивает возможность использования отдельных признаков для точного прогнозирования удоя или содержания жира у конкретного животного.

Для повышения прогностической ценности оценки экстерьера предлагается использовать не отдельные признаки, а их комбинации или комплексные индексы [12]. Такие комплексные показатели, объединяющие информацию о нескольких экстерьерных признаках, могут иметь более сильную связь с продуктивностью. Публикации в отечественной научной литературе свидетельствуют об активных исследованиях в этом направлении [13; 14].

В работе [15] был предложен подход к косвенной оценке уровней молочной продуктивности первотелок голштинской породы на основе комплексного балла линейных признаков оценки экстерьера (КБ). Авторы описали два алгоритма составления такого комплекса: первый основан на последовательном включении оценок признаков по убыванию их индивидуальной корреляции с общим показателем продуктивности, а второй – на выборе признаков, максимально увеличивающих коэффициент корреляции при добавлении к уже сформированному комплексу. В указанной работе было показано, что полиномиальные зависимости между рассчитанным КБ и средними значениями удоя и выхода молочного жира в группах животных характеризуются высокими коэффициентами детерминации ($R^2 > 0,95$) [15].

Несмотря на представленные высокие показатели аппроксимации, важным аспектом оценки практической применимости любой прогностической модели является ее валидация на независимых данных или с использованием строгих методов кросс-валидации [16].

Высокий R^2 при подгонке модели, особенно полиномиальной высокой степени, к ограниченному набору агрегированных данных (средних по группам с малым числом этих групп) может быть результатом переобучения [17] и не гарантирует аналогичной прогностической способности на новых данных. Кроме того, высокая аппроксимация средних значений по группам не равнозначна точной предсказательной способности для отдельных животных [18; 19].

Методология и методы исследования (Methods)

Оценка качества модели производилась для алгоритма составления КБ, рассчитываемого путем последовательного суммирования оценок линейных признаков в порядке убывания значений их

индивидуальной корреляции с общим показателем продуктивности [15].

Статистический анализ в рамках валидации предложенного метода не прямой оценки продуктивности на основе КБ проводился отдельно для двух разных типов данных для индивидуальных значений продуктивности каждого животного и для средних значений продуктивности групп животных, объединенных по значению КБ.

При оценке связи между КБ экстерьера и индивидуальными показателями молочной продуктивности, такими как удой за 305 дней лактации и выход молочного жира, был рассчитан ранговый коэффициент корреляции Спирмена (r_s). Этот непараметрический метод был выбран в связи с вероятным несоответствием распределения КБ критериям нормальности и использования в исходной методике его расчета преобразованных оценок признаков. Ранговая корреляция позволяет оценить силу и направление связи между двумя переменными, что является адекватным подходом при анализе связи между порядковыми или количественными данными имеющими распределение отличное от нормального [20].

Расчеты проводились отдельно для пары КБ и индивидуального удоя, а также для пары КБ и индивидуального выхода молочного жира.

Статистическая значимость полученных коэффициентов оценивалась стандартными методами при уровне значимости $\alpha < 0,05$. Интерпретация силы корреляционной связи выполнялась с опорой на общепринятые градации величины коэффициента корреляции [21].

Для оценки предсказательной способности КБ на уровне групп животных был применен многоэтапный статистический подход, воспроизводящий логику анализа, представленную в исходной работе [15].

Первым шагом было агрегирование данных: индивидуальные записи по каждой первотелке были сгруппированы на основе их КБ. Для каждой группы животных с одинаковым значением КБ были рассчитаны средние арифметические значения удоя и выхода молочного жира за 305 дней лактации. Одновременно определялась численность каждой группы (количество животных), поскольку репрезентативность средней оценки прямо зависит от объема данных, на которых она основана.

Следующим этапом стала фильтрация агрегированных данных. В соответствии с методикой исходной работы, из дальнейшего анализа были исключены те группы КБ, численность которых составила менее 10 наблюдений. Такой подход позволяет работать с более надежными оценками средних значений, минимизируя влияние выбросов и случайной вариабельности, особенно заметной в малочисленных группах [22]. В результате

для последующего моделирования использовался набор средних значений продуктивности, соответствующих только тем значениям КБ, для которых имелось достаточное количество индивидуальных наблюдений.

На наборе средних значений продуктивности по группам КБ были построены полиномиальные регрессионные модели 3-й степени. Выбор именно этой формы зависимости (полином 3-й степени) обусловлен тем, что, согласно исходной работе, она наилучшим образом аппроксимировала эмпирические зависимости средних показателей продуктивности от КБ. Были построены две отдельные модели: одна для прогнозирования среднего удоя группы по значению ее КБ, другая – для прогнозирования среднего выхода молочного жира. Формально модели описывались как: средний удой $\approx f(\text{КБ})$ и средний выход жира $\approx g(\text{КБ})$, где f и g являются полиномами третьей степени.

Для оценки того, насколько точно построенные на средних значениях полиномиальные модели будут предсказывать продуктивность для новых, ранее неиспользуемых групп КБ, была применен метод кросс-валидации (Leave-One-Out Cross-Validation, LOOCV) [23]. При LOOCV каждая из групп КБ последовательно исключалась из набора данных. Полиномиальная модель 3-й степени обучалась на оставшихся группах, а затем использовалась для предсказания среднего значения продуктивности для исключенной группы. Этот процесс повторялся для каждой группы, что позволило получить набор «кросс-валидированных» предсказаний для всех групп КБ, участвовавших в анализе. Качество этих предсказаний оценивалось с помощью стандартных метрик: среднеквадратическая ошибка (Root Mean Squared Error, RMSE), средняя абсолютная ошибка (Mean Absolute Error, MAE) и кросс-валидированный R-квадрат (R^2_{cv}). Эти показатели дали количественную оценку точности прогноза на уровне групп и доли объясненной вариации средних значений продуктивности на данных, не использованных непосредственно при подгонке модели [24; 25].

Для критической оценки статистической значимости результатов, полученных при аппроксимации средних значений молочной продуктивности полиномиальными моделями (что соответствует подходу, использованному в оригинальной работе [15]), был проведен Post-hoc-анализ статистической мощности. Целью данного анализа являлось определение вероятности того, что на основании F -теста, используемого для оценки общей значимости регрессионной модели, можно было бы обоснованно принять или отклонить нулевую гипотезу (отсутствие связи между КБ и средней продуктивностью) при наблюдаемой в исследовании силе эф-

фекта и имеющемся размере выборки ($n = 12$ агрегированных групп).

Все описанные статистические расчеты были выполнены в среде программирования R с использованием соответствующих пакетов для обработки данных.

Результаты (Results)

Анализ данных и валидация методов не прямой оценки молочной продуктивности на основе КБ проводились в два этапа: оценка связи КБ с индивидуальными показателями продуктивности и оценка прогностической способности моделей на агрегированных данных с использованием кросс-валидации.

На первом этапе была исследована взаимосвязь между КБ, представленным в исследуемой выборке ($n = 1243$), и индивидуальными значениями удоя и выхода молочного жира за 305 дней первой лактации. Был рассчитан ранговый коэффициент корреляции Спирмена (r_s). Ранговый коэффициент корреляции между КБ и индивидуальным удоём: $r_s = 0,287$, а между КБ и индивидуальным выходом молочного жира: $r_s = 0,230$.

Оба полученных коэффициента корреляции являются статистически значимыми при объеме выборки 1243 животных ($\alpha < 0,001$). Однако, согласно общепринятым шкалам интерпретации силы корреляции, значения r_s в диапазоне 0,2–0,3 считаются слабой связью [21]. Таким образом, несмотря на статистическую достоверность, КБ, использованный в данном исследовании, имеет ограниченную предсказательную способность для прогнозирования молочной продуктивности у отдельного животного.

Вариабельность удоя и выхода жира у индивидуальных коров в значительной степени определяется факторами, не отраженными в данном КБ.

На втором этапе для оценки возможности использования КБ для не прямой оценки продуктивности на уровне групп животных мы сгруппировали индивидуальные данные по значениям КБ и рассчитали средние показатели продуктивности для каждой группы. Всего в исходной выборке присутствовали животные с КБ в диапазоне от 36 до 55. После фильтрации, исключившей группы с численностью менее 10 наблюдений, для дальнейшего анализа были отобраны данные по 12 группам с КБ от 42 до 53. Средние значения удоя и выхода молочного жира представлены в таблице 1.

Анализируя таблицу 1, можно отметить явную тенденцию к увеличению среднего удоя и среднего выхода жира с ростом КБ, хотя строгой зависимости не наблюдается (например, средние значения для КБ = 46 немного ниже, чем для КБ = 45). Для описания такой зависимости в исходной работе использовались полиномиальные модели 3-й степени.

Таблица 1

Средние значения удоя и выхода молочного жира по группам КБ с численностью 10 и более голов

КБ	Численность группы первотелок, голов	Среднее значение	
		Удой, кг	Выход жира, кг
42	22	9 070,41	347,27
43	28	9 602,04	357,61
44	54	9 744,56	368,21
45	99	9 697,25	378,36
46	134	9 687,28	374,43
47	183	10 001,82	384,67
48	237	10 085,94	390,53
49	209	10 310,91	395,06
50	140	10 478,97	403,91
51	72	10 680,11	405,45
52	30	10 770,73	407,89
53	12	10 829,67	416,36

Table 1

Mean milk yield and milk fat yield by Complex Score (CS) groups with 10 or more first-calf heifers

CS	Number of heifers in the group, heads	Mean value	
		Milk yield, kg	Milk fat yield, kg
42	22	9 070.41	347.27
43	28	9 602.04	357.61
44	54	9 744.56	368.21
45	99	9 697.25	378.36
46	134	9 687.28	374.43
47	183	10 001.82	384.67
48	237	10 085.94	390.53
49	209	10 310.91	395.06
50	140	10 478.97	403.91
51	72	10 680.11	405.45
52	30	10 770.73	407.89
53	12	10 829.67	416.36

Полиномиальные регрессионные модели 3-й степени были построены на средних значениях, представленных в таблице 1. Анализ подгонки этих моделей ко всем 12 точкам показал очень высокие коэффициенты детерминации: R^2 для модели удоя составил 0,9538, а для модели выхода жира – 0,9845. Эти значения соответствуют данным, представленным в исходной работе [15], и указывают на то, что полином 3-й степени способен очень хорошо описывать наблюдаемую зависимость на имеющихся средних точках.

Для оценки прогностической способности моделей на уровне групп с использованием независимых данных была выполнена кросс-валидация на 12 группах (таблица 1). Результаты LOOCV представлены в таблице 2.

Результаты кросс-валидации показывают, что полиномиальные модели, построенные на средних значениях, обладают существенной прогностической способностью для предсказания продуктивности на уровне групп КБ. Для среднего удоя кросс-валидированный R^2 составил 0,7566. Это указывает на то, что модель способна объяснить около 76%

вариации среднего удоя между группами при проверке на независимых данных. Ошибки предсказания среднего удоя группы составляет около 178 кг (MAE) и 258 кг (RMSE). Для среднего выхода молочного жира прогностическая способность модели оказалась еще выше, R^2_{cv} для этой модели достиг 0,9688, т. е. полиномиальный тренд способен объяснить почти 97 % вариации среднего выхода жира между группами. Соответственно, ошибки предсказания среднего выхода жира группы очень низкие: MAE = 2,78 кг, RMSE = 3,61 кг.

Визуальное представление результатов LOOCV, где фактические средние значения по группам сравниваются с предсказанными моделью значениями (рис. 1 для удоя и рис. 2 для выхода жира), наглядно демонстрирует точность предсказаний на уровне групп. Точки на графиках располагаются близко к диагональной линии $y = x$, выполненной красным пунктиром. Особенно выражено это для выхода молочного жира, что подтверждает высокий R^2_{cv} и низкие ошибки предсказания на агрегированных данных.

Таблица 2

Метрики качества полиномиальных моделей 3-й степени, построенных на средних значениях продуктивности по группам КБ (LOOCV, $n = 12$ групп)

Модель для среднего показателя	RMSE (кг)	MAE (кг)	R^2_{cv}
Средний удой	257,87	177,84	0,7566
Средний выход жира	3,61	2,78	0,9688

Table 2

Quality metrics of the third-degree polynomial models fitted to mean productivity values by CS groups (LOOCV, $n = 12$ groups)

Model for the average indicator	RMSE (kg)	MAE (kg)	R^2_{cv}
Mean milk yield	257.87	177.84	0.7566
Mean milk fat yield	3.61	2.78	0.9688

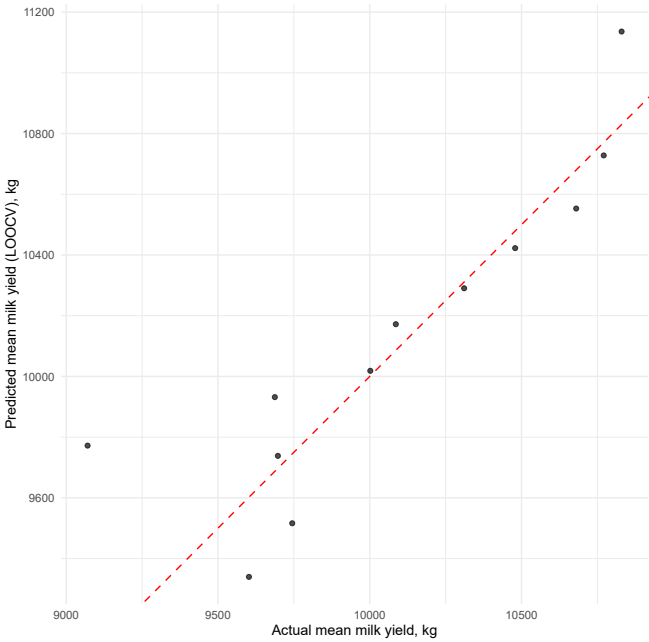


Рис. 1. Предсказанный и фактический средний удой ($R^2_{cv} = 0,7566$; $RMSE = 257,87$; $MAE = 177,84$)

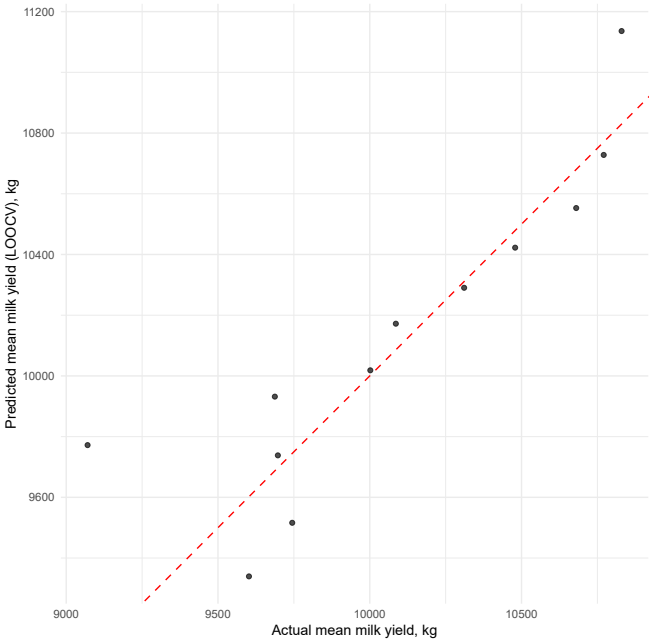


Fig. 1. Predicted vs actual mean milk yield from Leave-One-Out Cross-Validation ($R^2_{cv} > 0.7566$; $RMSE = 257.87$; $MAE = 177.84$)

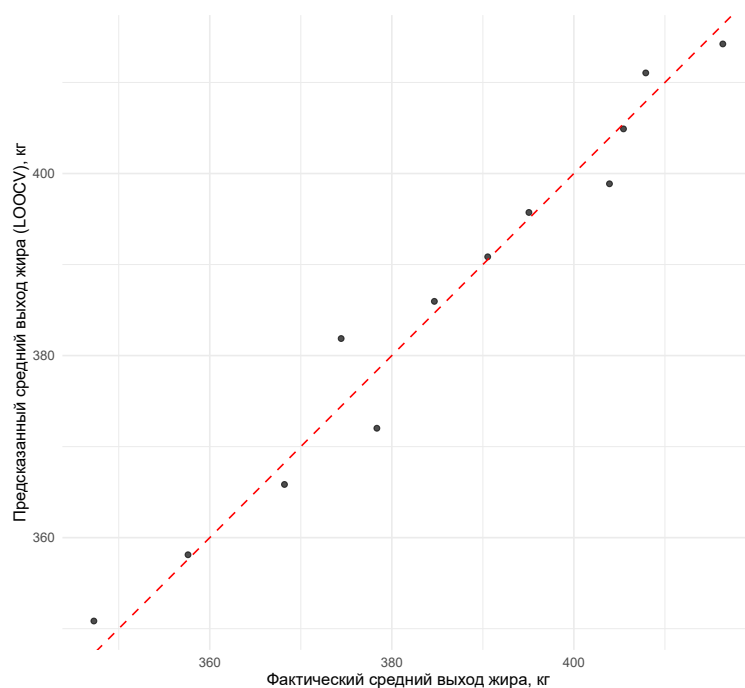


Рис. 2. Предсказанный и фактический средний выход жира ($R^2_{cv}=0,9688$; $RMSE=3,61$; $MAE=2,78$)

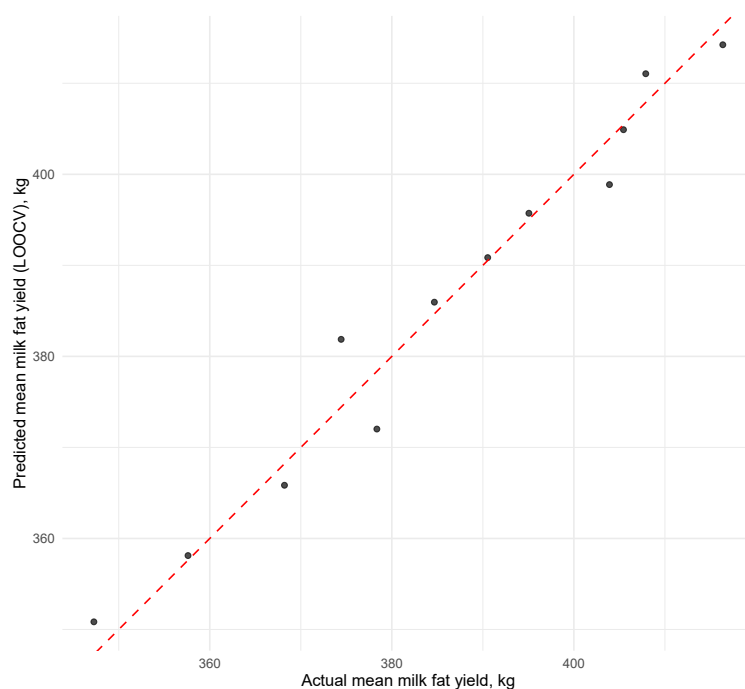


Fig. 2. Predicted vs actual mean milk fat yield from Leave-One-Out Cross-Validation ($R^2_{cv} > 0.9688$; $RMSE = 3.61$; $MAE = 2.78$)

Проведенный Post-hoc-анализ статистической мощности (таблица 3) показал, что при наблюдаемых размерах эффекта вероятность обнаружить статистически значимую связь на уровне $\alpha = 0,05$ составляла 100 %. Это согласуется с крайне низкими p -value ($p < 0,001$), полученными для F -теста моделей в целом.

Высокая расчетная мощность в данном контексте является прямым следствием очень сильной аппроксимации средних значений полиномом при небольшом количестве этих средних. Это свидетельствует о том, что наблюдаемые высокие значения R^2 (для средних по группам) статистически достоверно отличаются от нуля ($\alpha < 0,001$).

Таблица 3

Результаты Post-hoc-анализа статистической мощности полиномиальных моделей третьей степени для средних значений продуктивности ($n = 12$ групп)

Модель для среднего показателя	Наблюдаемый R^2	Наблюдаемый эффект f^2	Расчетная мощность (%)
Средний удой	0,9538	20,66	100,0
Средний выход жира	0,9845	63,53	100,0

Примечание. Анализ мощности проведен для F-теста модели при $\alpha = 0,05$ на 12 группах, при числителе степеней свободы $u = 3$ и знаменателе степеней свободы $v = 8$.

Table 3

Results of the Post-hoc power analysis for the third-degree polynomial models fitted to mean productivity values ($n = 12$ groups)

Model for the average indicator	Observed R^2	Observed effect f^2	Rated power (%)
Mean milk yield	0.9538	20.66	100.0
Mean milk fat yield	0.9845	63.53	100.0

Note. Power analysis was conducted for the model's F-test at $\alpha = 0.05$ on 12 groups, with numerator degrees of freedom $u = 3$ and denominator degrees of freedom $v = 8$.

Обсуждение и выводы (Discussion and Conclusion)

Проведенное исследование было направлено на валидацию первого метода непрямо́й оценки молочной продуктивности первотелок голштинской породы с использованием КБ, из двух предложенных в работе [15]. Результаты анализа позволяют всесторонне оценить предсказательную способность данного комплексного показателя, как на уровне индивидуальных животных, так и при работе с агрегированными данными по группам.

Анализ связи КБ с индивидуальными показателями удоя и выхода молочного жира выявил статистически значимую, но при этом слабую корреляцию ($r_s \approx 0,23 \dots 0,29$). Этот результат согласуется с многочисленными литературными данными, указывающими на низкую или умеренную взаимосвязь между отдельными признаками экстерьера и продуктивностью на уровне индивидуальных животных [26–28].

Несмотря на то что КБ, основанный на ряде линейных признаков, имеет высокую прогностическую способность для групп животных, его предсказательная возможность для точного прогнозирования продуктивности конкретной первотелки остается ограниченной. Это вполне объяснимо, так как на молочную продуктивность животного оказывает влияние огромное количество факторов, помимо морфологических особенностей экстерьера, включая генетический потенциал (не связанный с экстерьером), условия содержания, качество кормления, состояние здоровья, микроклимат и другие средовые воздействия. Эти многочисленные факторы создают значительную индивидуальную вариабельность продуктивности, которая существенно маскирует влияние экстерьерных признаков.

Однако, когда анализ был переведен на уровень агрегированных данных, мы получили иные результаты. Группировка животных по КБ и расчет сред-

них показателей продуктивности для достаточно многочисленных групп (с численностью не менее 10 голов) позволила сгладить значительную часть индивидуальной вариабельности, или «шума». На этом уровне была выявлена выраженная зависимость средних показателей продуктивности от КБ, которая, как показал анализ в исходной работе и подтвердило наше исследование, хорошо описывается полиномиальными трендами 3-й степени.

Результаты кросс-валидации наглядно продемонстрировали высокую прогностическую способность полиномиальных моделей для предсказания средних значений продуктивности групп, агрегированных на основании одинакового КБ. Высокие значения кросс-валидированного R^2 ($\approx 0,76$ для удоя и $\approx 0,97$ для выхода жира) указывают на то, что эти модели способны объяснять значительную долю вариации средних значений продуктивности между группами даже при оценке на данных, не использовавшихся напрямую при подгонке модели. Особенно впечатляющие результаты были получены для предсказания среднего выхода молочного жира, где модель показала практически полную объясняющую и прогностическую способность на уровне групп. Это говорит о том, что зависимость среднего выхода жира от КБ является более устойчивой и менее подверженной влиянию остаточной вариабельности даже после агрегации.

Таким образом, валидация подтверждает, что предложенный в исходной работе КБ и метод построения полиномиальных трендов на агрегированных данных могут быть эффективно использованы для непрямо́й оценки и прогнозирования среднего уровня молочной продуктивности в группах первотелок, объединенных по значению этого балла. Такая оценка на уровне групп может быть полезна в ряде практических сценариев: например, при оценке потенциала партий ремонтного молодняка, при

сравнении групп животных разного происхождения или выращенных в разных условиях, а также для принятия общих управленческих решений на уровне стада. Возможность косвенной оценки средней продуктивности групп становится особенно актуальной в контексте развития технологий автоматизированной и бесконтактной линейной оценки экстерьерера [29; 30], которые могут оперативно предоставлять данные для расчета КБ.

Ограничениями настоящего исследования являются использование данных, которые, вероятно, были частью выборки, использованной при разработке метода (что снижает степень независимости валидации), фокусировка только на первотелках голштинской породы и анализ продуктивности только за первую лактацию. Для более полной картины необходима валидация на полностью независимых выборках, включающих животных других пород, а также оценка взаимосвязи КБ с продуктивностью в

последующих лактациях и с показателями продолжительности хозяйственного использования.

Необходимо отметить, что, хотя в исходной работе оба метода расчета КБ демонстрировали схожие высокие R^2 при аппроксимации средних данных, второй алгоритм также нуждается в проведении валидации его прогностической способности.

Таким образом, проведенная валидация подтверждает пригодность предложенного в исходной работе КБ и соответствующего метода построения полиномиальных трендов для непрямой оценки и прогнозирования средней молочной продуктивности на уровне групп первотелок голштинской породы. Этот подход может служить полезным инструментом в селекционно-племенной работе и при принятии управленческих решений, позволяя проводить предварительную оценку потенциала групп животных, особенно в условиях массового использования автоматизированных систем линейной оценки экстерьерера.

Библиографический список

1. Коник Н. В., Калиниченко Э. Б., Каиров В. Р., Кубатиева З. А., Капитонова Е. А. Эффективность производственного использования коров молочных пород // Аграрный вестник Урала. 2025. Т. 25, № 01. С. 51–60. DOI: 10.32417/1997-4868-2025-25-01-51-60.
2. Татаркина Н. И., Свяженина М. А., Пономарева Е. А. Применение экстерьерной оценки в селекции крупного рогатого скота голштинской породы // Аграрный вестник Урала. 2023. Т. 23, № 10. С. 81–90. DOI: 10.32417/1997-4868-2023-23-10-81-90.
3. Романова Е. А., Тулинова О. В. Моделирование селекционного индекса для айрширской породы молочного скота с использованием экстерьерных показателей // Вестник Ульяновской государственной сельскохозяйственной академии. 2021. № 1 (53). С. 150–155. DOI: 10.18286/1816-4501-2021-1-150-155.
4. Девяткина Г. С., Молчанова Н. В., Сельцов В. И., Сулима Н. И. Линейная оценка коров черно-пестрой породы и ее связь с молочной продуктивностью // Вестник Российского университета дружбы народов. Серия: Агрономия и животноводство. 2010. № 2. С. 59–64.
5. Новиков А. В., Лешонок О. И. Взаимосвязь экстерьерера и молочной продуктивности коров-первотелок // Агропродовольственная политика России. 2014. № 4 (28). С. 49–51.
6. Батанов С. Д., Шайдуллина М. М. Продуктивные качества и экстерьерные особенности коров черно-пестрой породы разных линий // Ученые записки Казанской государственной академии ветеринарной медицины им. Н. Э. Баумана. 2019. Т. 239, № 3. С. 29–34. DOI: 10.31588/2413-4201-1883-239-3-29-35.
7. Новоселова К. С. Связь молочной продуктивности коров-первотелок с экстерьером в СХПК-СХА (колхоз) «Первое Мая» // Актуальные вопросы совершенствования технологии производства и переработки продукции сельского хозяйства. 2019. № 21. С. 356–359.
8. Madrid S., Echeverri J. Association between conformation traits and productive performance in Holstein cows in the department of Antioquia, Colombia // Veterinaria y Zootecnia. 2014. Vol. 8, No. 1. Pp. 35–47.
9. Batanov S., Baranova I., Starostina O. Innovative methods in study of animal's conformation // Bulgarian Journal of Agricultural Science. 2020. Vol. 26, No. 6. Pp. 1286–1291.
10. Misztal I., Lawlor T. J., Short T. H., VanRaden P. M. Multiple-trait estimation of variance components of yield and type traits using an animal model // Journal of Dairy Science. 1992. Vol. 75, No. 2. Pp. 544–551. DOI: 10.3168/jds.S0022-0302(92)77791-1.
11. Контэ А. Ф., Карликова Г. Г. Генетическая изменчивость показателей продуктивности и оценки экстерьерера голштинских коров в зависимости от типа телосложения // Аграрный вестник Урала. 2021. № 09 (212). С. 53–62. DOI: 10.32417/1997-4868-2021-212-09-53-62.
12. Toghiani S., VanRaden P. M. National index correlations and actual vs expected use of foreign sires // Leeuwarden. 2021. Vol. 56. Pp. 52–59.
13. Гарт В. В., Куликова С. Г., Богданова О. В., Норкина В. М., Камалдинов Е. В., Петров А. Ф. Полиномиальная сопряженная изменчивость признаков линейной оценки экстерьерера и удоя высокопродуктивного

голландского скота // Известия Тимирязевской сельскохозяйственной академии. 2024. № 5. С. 86–100. DOI: 10.26897/0021-342X-2024-5-86-100.

14. Гарт В. В., Куликова С. Г., Нарожных К. Н., Камалдинов Е. В. Раннее прогнозирование содержания молочного жира у голландского скота на основе сопряженной изменчивости с линейными признаками // Вестник НГАУ (Новосибирский государственный аграрный университет). 2024. № 4 (73). С. 168–176. DOI: 10.31677/2072-6724-2024-73-4-168-176.

15. Гарт В. В., Куликова С. Г., Камалдинов Е. В., Нарожных К. Н., Петров А. Ф. Непрямая оценка уровней признаков молочной продуктивности первотелок голландской породы // Аграрный вестник Урала. 2025. Т. 25, № 06. С. 900–914. DOI: 10.32417/1997-4868-2025-25-06-900-914.

16. Yates L. A., Aandahl Z., Richards S. A., Brook B. W. Cross validation for model selection: a review with examples from ecology // Ecological Monographs. 2023. Vol. 93, No. 1. Article number e1557. DOI: 10.48550/arXiv.2203.04552.

17. Hawkins D. M. The problem of overfitting // Journal of Chemical Information and Computer Sciences. 2004. Vol. 44, No. 1. DOI: 10.1021/ci0342472.

18. Olson K. M., Garrick D. J., Enns R. M. Predicting breeding values and accuracies from group in comparison to individual observations // Journal of Animal Science. 2006. Vol. 84, No. 1. Pp. 88–92.

19. Van der Vaart E., Johnston A. S. A., Sibly R. M. Predicting how many animals will be where: how to build, calibrate and evaluate individual-based models // Ecological Modelling. 2016. Vol. 326. Pp. 113–123.

20. Dikbaş F. A new two-dimensional rank correlation coefficient // Water Resources Management. 2018. Vol. 32. No. 5. Pp. 1539–1553.

21. Котеров А. Н., Ушенкова Л. Н., Зубенкова Э. С., Калинина М. В., Бирюков А. П., Ласточкина Е. М., Молодцова Д. В., Вайнсон А. А. Сила связи. Сообщение 2. Градации величины корреляции // Медицинская радиология и радиационная безопасность. 2019. Т. 64, № 6. С. 12–24. DOI: 10.12737/1024-6177-2019-64-6-12-24.

22. Aguinis H., Gottfredson R. K., Joo H. Best-practice recommendations for defining, identifying, and handling outliers // Organizational Research Methods. 2013. Vol. 16, No. 2. Pp. 270–301.

23. Wong T. T. Performance evaluation of classification algorithms by k-fold and leave-one-out cross validation // Pattern Recognition. 2015. Vol. 48, No. 9. Pp. 2839–2846.

24. Ostertagová E. Modelling using polynomial regression // Procedia Engineering. 2012. Vol. 48. Pp. 500–506.

25. Bruno D., Barca E., Goncalves R., Queiroz H., Berardi L., Passarella G. Linear and evolutionary polynomial regression models to forecast coastal dynamics: Comparison and reliability assessment // Geomorphology. 2018. Vol. 300. Pp. 128–140.

26. Madrid S., Echeverri J. Association between conformation traits and productive performance in Holstein cows in the department of Antioquia, Colombia // Veterinaria y Zootecnia. 2014. Vol. 8, No. 1. Pp. 35–47.

27. Batanov S., Baranova I., Starostina O. Innovative methods in study of animal's conformation // Bulgarian Journal of Agricultural Science. 2020. Vol. 26, No. 6. Pp. 1286–1291.

28. Misztal I., Lawlor T. J., Short T. H., VanRaden P. M. Multiple-trait estimation of variance components of yield and type traits using an animal model // Journal of Dairy Science. 1992. Vol. 75, No. 2. Pp. 544–551. DOI: 10.3168/jds.S0022-0302(92)77791-1.

29. Батанов С. Д., Баранова И. А., Старостина О. С. Разработка методов определения комплексного индекса типа телосложения с целью раннего прогнозирования молочной продуктивности и определения живой массы крупного рогатого скота // Известия Оренбургского государственного аграрного университета. 2023. № 1 (99). С. 299–307. DOI: 10.37670/2073-0853-2023-99-1-299-307.

30. Богданова О. В., Гарт В. В., Куликова С. Г., Камалдинов Е. В., Амерханов Х. А., Нарожных К. Н., Петров А. Ф., Жигулин Т. А., Астафьев А. А. Различия между странами по признакам линейной оценки экстерьера крупного рогатого скота голландской породы // Достижения науки и техники АПК. 2023. Т. 37, № 8. С. 59–64. DOI: 10.53859/02352451-2023-37-8-59.

Об авторах:

Владимир Владимирович Гарт, доктор сельскохозяйственных наук, профессор кафедры прикладной биоинформатики, Новосибирский государственный аграрный университет, Новосибирск, Россия; ORCID 0000-0002-7356-1090, AuthorID 92135. E-mail: gvlvl@yandex.ru

Светлана Геннадьевна Куликова, доктор биологических наук, профессор кафедры ветеринарной генетики и биотехнологии, Новосибирский государственный аграрный университет, Новосибирск, Россия; ORCID 0000-0002-0341-5055, AuthorID 96363. E-mail: kulikovasg@yandex.ru

Евгений Варисович Камалдинов, доктор биологических наук, заведующий кафедрой прикладной биоинформатики, Новосибирский государственный аграрный университет, Новосибирск, Россия; ORCID 0000-0002-0341-5055, AuthorID 336282. E-mail: ekamaldinov@yandex.ru

Алексей Федорович Петров, заведующий лабораторией прикладной биоинформатики, Новосибирский государственный аграрный университет, Новосибирск, Россия; ORCID 0000-0002-7402-4107, AuthorID 1071714. E-mail: alex@nsau.edu.ru

Кирилл Николаевич Нарожных, кандидат биологических наук, доцент кафедры прикладной биоинформатики, Новосибирский государственный аграрный университет, Новосибирск, Россия; ORCID 0000-0002-1519-697X, AuthorID 654429. E-mail: narozhnykh@nsau.edu.ru

References

1. Konik N. V., Kalinichenko E. B., Kairov V. R., Kubatieva Z. A., Kapitonova E. A. Efficiency of production use of dairy cows in the context of federal districts of Russia. *Agrarian Bulletin of the Urals*. 2025; 25 (01): 51–60. DOI: 10.32417/1997-4868-2025-25-01-51-60. (In Russ.)
2. Tatarkina N. I., Svyazhenina M. A., Ponomareva E. A. The use of exterior assessment in the selection of Holstein cattle. *Agrarian Bulletin of the Urals*. 2023; 23 (10): 81–90. DOI: 10.32417/1997-4868-2023-23-10-81-90. (In Russ.)
3. Romanova E. A., Tulinova O. V. Breeding index modeling for airshire dairy cattle with application of exterior parameters. *Vestnik of Ulyanovsk State Agricultural Academy*. 2021; 1 (53): 150–155. DOI: 10.18286/1816-4501-2021-1-150-155. (In Russ.)
4. Devyatkina G. S., Molchanova N. V., Seltsov V. I., Sulima N. I. Linear estimation of cows of black-and-white breed and its correlation with dairy production. *RUDN Journal of Agronomy and Animal Industries*. 2010; 2: 59–64.
5. Novikov A. V., Leshonok O. I. Relationship of the exterior and milk production of cows-heifers. *Russian Agricultural and Food Policy*. 2014; 4: 49–51.
6. Batanov S. D., Shaydullina M. M. Productive qualities and exterior features of black-and-white breeds cows of different lines. *Scientific Notes Kazan Bauman State Academy of Veterinary Medicine*. 2019; 2399 (3): 29–34. DOI: 10.31588/2413-4201-1883-239-3-29-35 (In Russ.)
7. Novoselova K. S. The relationship of dairy productivity of first-time cows with the exterior in SKHPK-SKHA (kolkhoz) “Pervoye Maya”. *Current Issues of Improving the Technology of Production and Processing of Agricultural Products*. 2019; 21: 356–359. (In Russ.)
8. Madrid S., Echeverri J. Association between conformation traits and productive performance in Holstein cows in the department of Antioquia, Colombia. *Veterinaria y Zootecnia*. 2014; 8 (1): 35–47.
9. Batanov S., Baranova I., Starostina O. Innovative methods in study of animal's conformation. *Bulgarian Journal of Agricultural Science*. 2020; 26 (6): 1286–1291.
10. Misztal I., Lawlor T. J., Short T. H., VanRaden P. M. Multiple-trait estimation of variance components of yield and type traits using an animal model. *Journal of Dairy Science*. 1992; 75 (2): 544–551. DOI: 10.3168/jds.S0022-0302(92)77791-1.
11. Konte A. F., Karlikova G. G. Genetic variability of productivity traits and evaluation of exterior of Holstein cows depending on body type. *Agrarian Bulletin of the Urals*. 2021; 09 (212): 53–62. DOI: 10.32417/1997-4868-2021-212-09-53-62. (In Russ.)
12. Toghiani S., VanRaden P. M. National index correlations and actual vs. expected use of foreign sires. *Leewarden*. 2021; 56: 52–59.
13. Gart V. V., Kulikova S. G., Bogdanova O. V., Norkina V. M., Kamaldinov E. V., Petrov A. F. Polynomial conjugate variability of traits of linear assessment of the exterior and milk yield of highly productive Holstein cattle. *Izvestiya of Timiryazev Agricultural Academy*. 2024; 5: 86–100. DOI: 10.26897/0021-342X-2024-5-86-100. (In Russ.)
14. Gart V. V., Kulikova S. G., Narozhnykh K.N., Kamaldinov E.V. Early prediction of milk fat content in Holstein cattle based on correlated variability with linear traits. *Bulletin of NSAU (Novosibirsk State Agrarian University)*. 2024; 4 (73): 168–176. DOI: 10.31677/2072-6724-2024-73-4-168-1764. (In Russ.)
15. Gart V. V., Kulikova S. G., Kamaldinov E. V., Narozhnykh K. N., Petrov A. F. Indirect assessment of the levels of milk productivity traits in first-calf heifers of the Holstein breed. *Agrarian Bulletin of the Urals*. 2025; 25 (06): 900–914. DOI: 10.32417/1997-4868-2025-25-06-900-914. (In Russ.)
16. Yates L. A., Aandahl Z., Richards S. A., Brook B. W. Cross validation for model selection: a review with examples from ecology. *Ecological Monographs*. 2023; 93 (1): e1557. DOI: 10.48550/arXiv.2203.04552.
17. Hawkins D. M. The problem of overfitting. *Journal of Chemical Information and Computer Sciences*. 2004; 44 (1). DOI: 10.1021/ci0342472.

18. Olson K. M., Garrick D. J., Enns R. M. Predicting breeding values and accuracies from group in comparison to individual observations. *Journal of Animal Science*. 2006; 84 (1): 88–92.
19. Van der Vaart E., Johnston A. S. A., Sibly R. M. Predicting how many animals will be where: how to build, calibrate and evaluate individual-based models. *Ecological Modelling*. 2016; 326: 113–123.
20. Dikbaş F. A new two-dimensional rank correlation coefficient. *Water Resources Management*. 2018; 32 (5): 1539–1553.
21. Koterov A. N., Ushenkova L. N., Zubenkova E. S., Kalinina M. V., Biryukov A. P., Lastochkina E. M., Molodtsova D. V., Wainson A. A. Strength of Association. Report 2. Graduations of Relative Risk. *Medical Radiology and Radiation Safety*. 2019; 64 (6): 12–24. DOI: 10.12737/1024-6177-2019-64-6-12-24.
22. Aguinis H., Gottfredson R. K., Joo H. Best-practice recommendations for defining, identifying, and handling outliers. *Organizational Research Methods*. 2013; 16 (2): 270–301.
23. Wong T. T. Performance evaluation of classification algorithms by k-fold and leave-one-out cross validation. *Pattern Recognition*. 2015; 48 (9): 2839–2846.
24. Ostertagová E. Modelling using polynomial regression. *Procedia Engineering*. 2012; 48: 500–506.
25. Bruno D., Barca E., Goncalves R., Queiroz H., Berardi L., Passarella G. Linear and evolutionary polynomial regression models to forecast coastal dynamics: Comparison and reliability assessment. *Geomorphology*. 2018; 300: 128–140.
26. Madrid S., Echeverri J. Association between conformation traits and productive performance in Holstein cows in the department of Antioquia, Colombia. *Veterinaria y Zootecnia*. 2014; 8 (1): 35–47.
27. Batanov S., Baranova I., Starostina O. Innovative methods in study of animal's conformation. *Bulgarian Journal of Agricultural Science*. 2020; 26 (6): 1286–1291.
28. Misztal I., Lawlor T. J., Short T. H., VanRaden P. M. Multiple-trait estimation of variance components of yield and type traits using an animal model. *Journal of Dairy Science*. 1992; 75 (2): 544–551. DOI: 10.3168/jds.S0022-0302(92)77791-1.
29. Batanov S. D., Baranova I. A., Starostina O. S. Development of methods for determining a complex index of body type for the purpose of early forecasting of milk production and determining the live weight of cattle. *Izvestia Orenburg State Agrarian University*. 2023; 1 (99): 299–307. DOI: 10.37670/2073-0853-2023-99-1-299-307. (In Russ.)
30. Bogdanova O. V., Gart V. V., Kulikova S. G., Kamaldinov E. V., Amerkhanov Kh. A., Narozhnykh K. N., Petrov A. F., Zhigulin T. A., Astafyev A. A. Differences between countries in terms of linear assessment of the conformation of Holstein cattle. *Achievements of Science and Technology of AIC*. 2023; 37 (8): 59–64. DOI: 10.53859/02352451_2023_37_8_59. (In Russ.)

Authors' information:

Vladimir V. Gart, doctor of agricultural sciences, professor of applied bioinformatics chair, Novosibirsk State Agrarian University, Novosibirsk, Russia; ORCID 0000-0002-7356-1090, AuthorID 92135.

E-mail: gvlvl@yandex.ru

Svetlana G. Kulikova, doctor of biological sciences, professor of veterinary genetics and biotechnology chair, Novosibirsk State Agrarian University, Novosibirsk, Russia; ORCID 0000-0002-0341-5055, AuthorID 96363,

E-mail: kulikovasg@yandex.ru

Evgeniy V. Kamaldinov, doctor of biological sciences, head of applied bioinformatics chair, Novosibirsk State Agrarian University, Novosibirsk, Russia; ORCID 0000-0002-0341-5055, AuthorID 336282.

E-mail: ekamaldinov@yandex.ru

Aleksey F. Petrov, head of applied bioinformatics laboratory, Novosibirsk State Agrarian University, Novosibirsk, Russia; ORCID 0000-0002-7402-4107, AuthorID 1071714. E-mail: alex@nsau.edu.ru

Kirill N. Narozhnykh, candidate of biological sciences, associate professor of applied bioinformatics chair, Novosibirsk State Agrarian University, Novosibirsk, Russia; ORCID 0000-0002-1519-697X, AuthorID 654429,

E-mail: narozhnykh@nsau.edu.ru