Impact of different numbers of milk test-day records during lactation on the reliability of estimated breeding values

Giovani Luis Feltes1 · Vanessa Tomazetti Michelotti1 · Mauricio Morgado Oliveira1 · Fernanda Cristina Breda1 · Paulo Roberto Nogara Rorato1 · Renata Negri2 · Alessandro Haiduck Padilha2 · Jaime Araújo Cobuci2

Received: 3 March 2022 / Accepted: 31 August 2022 / Published online: 15 September 2022 © The Author(s), under exclusive licence to Springer Nature B.V. 2022

Abstract
The objective of this study was to evaluate the impact of the reduction in the number of test-day records per lactation on genetic parameters of test-day milk yield (TDMY) and the reliability of estimated breeding values for 305-day milk yield in Holstein cattle. Estimates of genetic parameters and breeding values were performed using the animal model of random regression and adjustment of the Legendre polynomial (fourth order). When comparing sires with the same number of daughters, greater reliability was found in the subpopulations with the highest number of milk test-day records per lactation per cow. It was also found that the elimination of at least one test-day record affected the reliability of estimated breeding value for 305-day milk yield in the sires, regardless of the class of number of daughters per sire. When selecting the 5% best sires and 20% best cows, the lowest order correlations were observed between the population with 10 test-day records per lactation (complete lactation) and the other subpopulations with incomplete lactations (4, 5, 6, 7, 8, 9 test-day records per lactation). The reduction in the number of test-day milk records per lactation interferes in the reliability of the estimated breeding value for sires and cows, negatively impacting the precision of selecting genetically superior animals.

Keywords Genetic evaluation · Spearman’s correlation · Selection · Random regression

Introduction
The use of the complete lactation curve to genetically select animals has been a complex task in recent decades, and to simplify the process of genetic evaluation in breeding programs worldwide, the feasibility of selecting specific patterns of lactation curves or derived traits, such as persistence in lactation and peak lactation, has been extensively investigated (Oliveira et al., 2019a). It is observed that a series of studies started to use the random regression models and, later, included genomic data in these models to estimate parameters and estimated breeding values more accurately in relation to previous decades (Oliveira et al., 2019a, 2019b).

The use of random regression models in the genetic evaluation of dairy cattle is widespread and has been adopted by several countries in their national genetic evaluations (Dzomba et al., 2011). In these evaluations, the models were adjusted to the yield records collected monthly during lactation by technicians belonging to the milk control service of breeders’ associations or private companies, generating about 10 test-day milk yield records in a lactation period of 305 days.

However, not all herds or all animals are controlled monthly due to the costs of carrying out milk control (Padilha et al., 2017) and the lack of performance and genetic records in certain countries (Ojango et al., 2016). Thus, a significant part of the cows present in the genetic evaluation may have less than 10 milk test-day records per lactation. The local costs for carrying out the milk control are related to the transportation, food, and accommodation of the professional who performs the control, who must be paid by the breeder, which can vary between 6.6 and 17.0% of the net revenue of the property (Cardoso et al., 2005).
Due to the high cost (where there is no type of government subsidy), some Breeder Associations are recommending a decrease in the frequency of milk test-day records per lactation per cow (Rural Pecuária, 2016); however, with the maintenance of this indication, the average number of test-days per cows should decrease and this could lead to significant changes in genetic evaluation. Therefore, it is necessary to carry out studies to assess the possible impacts.

In this context, it is common to find several studies in the literature with random regression models that used different requirements in the database in relation to the minimum number of test days per lactation per cow (3, 4, 5, 6, and 8), such as those carried out in Iran (Torshizi et al., 2013; Mohammadi et al., 2014; Naserkheil et al., 2016), Brazil (Biassus et al., 2011), and Ethiopia (Meseret and Negussie, 2017). This confirms that not all cows in genetic evaluation have 10 milk test-days per lactation.

On the other hand, justifications for choosing these minimum numbers of test-day records are not commonly reported. However, the authors mentioned above concluded that with a certain minimum number of test-day records per cow per lactation, it is possible to estimate the genetic parameters of a population and estimated breeding values of the animals, but they did not evaluate in detail the reflexes of this type of restriction in the records (or imposition of the data sets themselves) on the classification of animals and on impacts of the accuracy of the selection process. Nevertheless, the question arises as to what minimum number of test-day records per cow per lactation are necessary not to impose bias on the genetic evaluation of populations where there are limitations non availability of 10 test-day milk records per cow.

The objective in this work was to evaluate the impact of the reduction in the number of milk test-day records per lactation per cow on the quality of genetic evaluation of Holstein sires and cows managed in different Brazilian production systems.

### Material and methods

The original data file used in this study is described in Padelha et al. (2017). The original data were restricted to the TDMY of Holstein cows aged between 18 and 48 months and included cows of first lactation. This same database has undergone a new edition through software R (2017) to meet the demands of this work and still obtain data consistency. Restrictions were made to keep only animals with 10 lactation test-day records, one per month, between 6 and 305 days of lactation.

The contemporary groups were defined by herd-year-month (HYM) of test-day records, excluding contemporary groups with less than four animals, as well as TDMY records below or above 3.5 standard deviations from the mean to ensure better data consistency. Sires with daughters in only one herd were also excluded.

Following these criteria, a total of 111,840 TDMY records from 11,184 first lactations of Holstein cows of 128 herds, daughters of 1,233 sires, were analyzed. The population formed with all cows with 10 test-day records was called P10. Six other subpopulations were also named based on the random reduction of test-day records: reduction of six records (P4), reduction of five records (P5), reduction of four records (P6), reduction of three records (P7), reduction of two records (P8), and reduction of one record (P9). More details on subpopulations are available in Table 1.

For each lactating cow, a number from 1 to 10 was drawn, and the milk test-day records corresponding to that number was eliminated; after this process, again the number of animals per contemporary group (restricting to maintain a minimum of four animals in all subpopulations) was verified and with this the P9 population was formed. To form the P8 population, for each cow, two numbers from 1 to 10 were drawn, and the milk test-day records corresponding to those numbers were eliminated, and so on until P4, in which, in the formation of this last subpopulation, six numbers from

### Table 1 Description of the subpopulations of first lactating Holstein cows raised in Brazil

| Items          | P10 | P9  | P8  | P7  | P6  | P5  | P4  |
|----------------|-----|-----|-----|-----|-----|-----|-----|
| Cows           | 11,184 | 9743 | 8692 | 7773 | 6655 | 5912 | 4995 |
| Sires          | 1233  | 1160 | 1098 | 1051 | 984  | 937  | 866  |
| Pedigree file  | 24,090 | 21,284 | 19,188 | 17,675 | 14,032 | 14,032 | 12,139 |
| HYM            | 6142  | 4686 | 3863 | 3214 | 2530 | 2143 | 1556 |
| Herds          | 128   | 93   | 81   | 64   | 53   | 45   | 35   |
| TDMY           | 27.32 | 27.62 | 27.83 | 27.97 | 28.15 | 28.26 | 28.55 |
| AFC            | 26.44 | 26.27 | 26.17 | 26.1 | 25.88 | 25.77 | 25.56 |

P10, 10 records per lactation; P9, nine records per lactation; P8, eight records per lactation; P7, seven records per lactation; P6, six records per lactation; P5, five records per lactation; P4, four records per lactation; HYM, contemporary groups; TDMY, average test-day milk yield (kg); AFC, age at first calving (months)
1 to 10 were drawn and the records related to these numbers were eliminated.

The genetic parameters and breeding values were estimated using a random regression model and adjustment of the Legendre orthogonal polynomial (fourth order), as:

\[ Y_{ijklm} = HYM_j + \sum_{i=1}^{n} a_i \Theta_i(DIM) + \hat{\beta}_{mk} \]

\[ + \sum_{t=0}^{m} \Theta_t(DIM) \gamma_t + \sum_{t=0}^{m} \Theta_t(DIM) \gamma_t \]

\[ + \epsilon_{ijklm}, \]

in which \( Y_{ijklm} \) is the test-day records performed in month \( t \) within the herd-year-month of test-day record \( HYM_j \); \( HYM_j \) is the fixed effect of contemporary group (herd-year-month of calving); \( \hat{\beta}_{mk} \) is the fixed regression coefficient, defined as the age classes 1 (18 to 25 months), 2 (26 to 27 months), 3 (28 to 29 months), and 4 (30 to 48 months), combined with the calving season subclasses 1 (January to March), 2 (April to June), 3 (July to September), and 4 (October to December), totaling sixteen fixed curves; \( u_{jk} \) is the \( j \)th random regression coefficient for the additive genetic effect of the \( j \)th cow; \( \gamma_t \) is the Legendre polynomial for test-day month \( t \) of the \( j \)th cow, in which \( k \) is the Legendre polynomial coefficient; and \( \epsilon_{ijklm} \) is the random residual.

We assume: \( \text{Var} \begin{bmatrix} u & \text{A} \otimes G & G & 0 & 0 \\ pe & 0 & I \otimes P & 0 \\ e & 0 & 0 & R \end{bmatrix} \), in which \( G \) and \( P \) are the covariance matrices of the random regression coefficients, \( A \) is the relationship matrix, \( I \) is the identity matrix, \( R = \sigma^2_e \) is a diagonal matrix (residual), and \( \otimes \) is the Kronecker product.

To adjust the Legendre polynomials, we first proceeded to standardize the test day milk control (t) to have magnitudes between –1 and 1, as proposed by Kirkpatrick et al. (1994), \( dim_r = -1 + \frac{2(dim_{\text{min}}-dim_r)}{dim_{\text{max}}-dim_{\text{min}}} \), in which \( dim_{\text{min}} \) and \( dim_{\text{max}} \) are minimum and maximum values for test day milk control; \( t^* \) is the standardized record day (\( dim_r \)); the polynomial \( k \) is given as follows:

\[ \Theta(dim_r,k) = \frac{1}{\sqrt{\pi}} \sum_{n=0}^{k/2} \frac{(k/2)!}{n!(k/2-n)!} \left( \frac{2}{\pi} \right)^{k/2} \left( \frac{2}{\pi} \right)^{(2k+1)/2} \times, \]

in which \( K/2 = (k-1)/2, \) if \( k \) is odd.

The estimated breeding value of animal \( i \) in month \( t \) was calculated as \( EBV_{g_i} = \hat{z}_i \hat{a}_i = \sum_{t=0}^{k-1} a_i \Theta_t(dim_r) \), being \( \hat{a}_i \) a vector \( (k_x 1) \) of the estimates of the additive genetic random regression coefficients of the animal \( i \), and \( \hat{z}_i \) a vector of Legendre polynomial coefficients for month \( t \). The genetic analysis were made using the software REMFL90 (Misztal et al., 2002) that uses the maximum likelihood restricted method.

The reliability values were obtained by \( r^2 = 1 - (SE^2/\sigma^2_a) \), in which \( SE \) is the standard error, which is the square root of the variance of the breeding values prediction error, \( \sigma^2_a \) is the additive genetic variance for the trait, and \( r^2 \) is the correlation between the true and the estimated breeding value (Misztal and Wiggans, 1988).

To compare the results obtained for the different subpopulations, the average reliability of the breeding value for 305-day milk yield was calculated for the sire of bulls (SB5 = 5% best sires), sire of cows (SC20 = 20% best sires), dams of bulls (DB2 = 2% best cows), and dams of cows (DC85 = 85% best cows). We sought to group the sires in terms of their number of daughters in six classes: class A = sires with up to 10, B = 11 to 30, C = 31 to 50, D = 51 to 75, E = 76 to 100, and F = more than 100.

Subsequently, we calculated Spearman’s ranking correlation and coincidence of the animals in the seven subpopulations in the different ranking intensities (SB5, SC20, DB2, and SC85) for estimated breeding values for 305-day milk yield.

Results

Variance components and genetic parameters

The estimated residual variance did not show evident differences between the subpopulations, with the lowest estimate being obtained for the subpopulations P9 (5.19 kg²) and the highest for P4 (5.58 kg²). When considering the estimates obtained in the analyses of the subpopulations P10 (5.31 kg²), P8 (5.32 kg²), P7 (5.30 kg²), P6 (5.53 kg²), and P5 (5.47 kg²) together with the two previously mentioned, an average value of 5.38 kg² could be verified with a standard deviation of 0.14 kg², indicating almost no change in the residual variance estimates between subpopulations.

The estimated permanent environment variances ranged from 7.46 kg² (month 8, P10) to 18.23 kg² (month 1, P5), with similar behavior throughout lactation among the seven subpopulations evaluated (Fig. 1), presenting a “U” shaped curve, that is, with the highest estimates observed for the beginning and end of lactation. There was a small difference between the values when comparing the estimates obtained in P10 with those observed in the other subpopulations.

The fact that there is no information on complete lactation led to an underestimation of the values of genetic variance in the other subpopulations (P4 to P9). These in turn showed small differences in values between themselves for the lactation months. The highest estimates of genetic variance were observed for P10, which ranged from 6.94 to 15.09 kg². In the other subpopulations, these estimates ranged from 3.20 to 8.30 kg² (Fig. 1).

The differences in values of genetic and phenotypic variance estimates observed between P10 and other
subpopulations made it possible to obtain higher heritability values throughout lactation at P10, which ranged from 0.25 to 0.52 (Fig. 1). In the other subpopulations (P4 to P9), the heritability estimates were close to each other and varied from 0.14 to 0.34 (P9 = 0.14 to 0.31, P8 = 0.14 to 0.33, P7 = 0.17 to 0.31, P6 = 0.17 to 0.31, P5 = 0.15 to 0.34, and P4 = 0.20 to 0.30) over the lactation period.

The decrease in the number of test-day record did not affect the correlation of permanent environment, because all the subpopulations had an identical format (data not shown), varying from 0.19 to 0.98, with the highest values among test-day record closest to each other.

Differently from what was observed for permanent environment correlation, the genetic correlation between lactation months differed between P10 and the other subpopulations. At P10, the estimate of genetic correlation between milk yields in different lactation months ranged from 0.75 and 0.99, with greater correlation between adjacent months and smaller between those distant.

In the other subpopulations, the genetic correlation varied more, for example P9 (0.27 to 0.99) and P5 (0.47 to 0.99), but with the same behavior of higher correlations between milk yield in adjacent months and lower among milk yield in distant months.

**Reliability of estimated breeding value**

Regardless of the number of daughters per sire, the average reliability was higher in the subpopulation with the highest number of milk test-day records per lactation (P10). The average reliability of the estimated breeding value for 305-day milk yield (Table 2) of SB5 varied from 0.46 (P4) to 0.74 (P10) and of SC20, from 0.36 (P4) to 0.64 (P10). For females, the average reliability of DB2 and DC85 ranged from 0.40 (P4) to 0.69 (P10), and from 0.23 (P4) to 0.45 (P10), respectively.

The absence of only one record is sufficient to cause a decrease in the average reliability of the estimated breeding value (Fig. 2). In the comparison among P9, P8, P7, P6, P5, and P4, the average reliability was similar, with a slight tendency to decrease as the number of test-day records decreased.

The number of sires present in the genetic evaluation of these subpopulations with fewer records reduced drastically, especially sires with 50 or more daughters, classes D, E, and F (Table 3), which are the sires with the greatest reliability of their breeding value: in class D, reduction of 40 sires in P10 to 19 sires in P4; in class E, reduction of 20 sires in P10 to four sires in P4; and in class F, reduction of 17 sires in P10 to one sire in P4.
For all populations, sires with a greater number of daughters had more reliable estimated breeding value; however, the reduction in the number of records also reduced the number of daughters evaluated for the sires, thus implying a reduction in the mean reliability between populations P10 and P4 of 28, 28, 29, and 22% in S2, S20, C2, and C85, respectively.

**Spearman’s ranking correlation**

Considering the different selection intensities, both in males and females, it was found that the Spearman’s correlations decreased with the increase in the reduction of milk test-day records per cow per lactation (Fig. 3).

| Subpopulation | SB5 | SC20 | DB2 | DC85 |
|---------------|-----|------|-----|------|
| P10           | 0.74 (0.40 to 0.95) | 0.64 (0.08 to 0.97) | 0.69 (0.32 to 0.75) | 0.45 (0.00 to 0.77) |
| P9            | 0.59 (0.24 to 0.89) | 0.45 (0.05 to 0.89) | 0.45 (0.21 to 0.55) | 0.29 (0.00 to 0.57) |
| P8            | 0.57 (0.17 to 0.89) | 0.45 (0.05 to 0.89) | 0.46 (0.19 to 0.56) | 0.29 (0.00 to 0.57) |
| P7            | 0.55 (0.16 to 0.87) | 0.43 (0.04 to 0.88) | 0.44 (0.17 to 0.54) | 0.27 (0.00 to 0.55) |
| P6            | 0.55 (0.15 to 0.86) | 0.42 (0.04 to 0.86) | 0.42 (0.16 to 0.52) | 0.26 (0.00 to 0.53) |
| P5            | 0.55 (0.16 to 0.86) | 0.42 (0.04 to 0.86) | 0.44 (0.21 to 0.53) | 0.27 (0.00 to 0.55) |
| P4            | 0.46 (0.14 to 0.84) | 0.36 (0.02 to 0.84) | 0.40 (0.19 to 0.50) | 0.23 (0.00 to 0.51) |

**Fig. 2** Average reliability of the estimated breeding value for 305-day milk yield of the sires with different number of daughters evaluated in the subpopulations (P10 = 10 records per lactation; P9 = nine records per lactation; …; P4 = four records per lactation)

**Table 3** Number of sires with different number of daughters (class A = sires with up to 10 daughters; B = 11 to 30; C = 31 to 50; D = 51 to 75; E = 76 to 100; and F = more than 100) evaluated in the subpopulations (P10 = 10 records per lactation; P9 = nine records per lactation; …; P4 = four records per lactation)

| Class | P10 | P9 | P8 | P7 | P6 | P5 | P4 |
|-------|-----|----|----|----|----|----|----|
| A     | 815 | 783| 752| 728| 701| 675| 630|
| B     | 269 | 245| 224| 211| 186| 179| 164|
| C     | 72  | 67 | 64 | 65 | 55 | 52 | 48 |
| D     | 40  | 35 | 33 | 28 | 28 | 20 | 19 |
| E     | 20  | 20 | 16 | 13 | 11 | 9  | 4  |
| F     | 17  | 10 | 9  | 6  | 3  | 2  | 1  |
The lowest Spearman’s correlations were observed between P10 and the other subpopulations, ranging from 0.33 (P10 vs P7) to 0.55 (P10 vs P9) in the case of S5, and for S20, Spearman’s correlations ranged from 0.46 (P10 vs P4) to 0.64 (P10 vs P9). In females, Spearman’s correlations for C2 ranged from 0.31 (P10 vs P6) to 0.37 (P10 vs P8), and for C85, from 0.6 (P10 vs P4) to 0.73 (P10 vs P9). In general, populations with close number of test-day records have higher ranking correlation values.

Greater coincidences (Fig. 4) were reported in the selection of C85, all greater than 90%; for the other selectable groups, the coincidence was lower, but always above 50%. The smallest coincidences were always between the selected groups at P10 compared with the other subpopulations (P9 to P4).

Discussion

Variance components and genetic parameters

The residual variance estimated in the different subpopulations did not vary much between them, and the values are similar to those previously reported for the first lactation of Holstein animals in Brazil by Cobuci et al. (2011), Araújo et al. (2006), and Costa et al. (2008), and are low values, which according to Damane et al., (2016), are indicative of a good fit of the model; with data from different populations, it can be said that the residual variance has not changed.

The permanent environment variance presented a format already reported for the Holstein breed in Brazil (Bignardi et al., 2009; Cobuci et al., 2011), and this format can be attributed to the beginning of lactation, with the adaptation of the first animals to milking and also to the negative energy balance, and at the end of lactation, to the different methods of drying the animals via changes in feeding (Damane et al., 2016). These differences were also attributed to the difficulty of modeling and the lower number of test-day records at the beginning and at the end of lactation by Bignardi et al. (2009).

The variance of permanent environment in P10 was slightly less than that of other populations in all months, so the reduction in the number of records slightly altered the value of this variance but did not change its behavior throughout the lactation period.

The shape of the genetic variance curve, with lower values at the beginning of lactation and increasing over the months, was the same for all populations; however, this increase was greater in P10 compared with the others, thus being able to affirm that the genetic variance is affected by the number of records. The values obtained for P10 agree with studies that used population with complete lactation records (10 monthly test-day records) as the studies of Olori et al. (1999) in the United Kingdom, in Multiple Ovulation and Embryo Transfer (MOET), and Zavadilová et al. (2005) in the Czech Republic.

The heritability of milk production in P10 was higher than in the other subpopulations, values similar to those reported by Olori et al. (1999) working with weekly test-day records.

Fig. 3 Spearman’s ranking correlation between sires of bulls (above the diagonal) and sires of cows (below the diagonal) and between dams of bulls (above the diagonal) and dams of cows (below the diagonal) based on the breeding value for 305-day milk yield in the subpopulations (P10 = 10 records per lactation; P9 = nine records per lactation; …; P4 = four records per lactation)
records, and Zavadilová et al. (2005) working with an average above 9.3 test-day records per cow per lactation. For the other populations (P4 to P9), heritability is similar to the studies of Araújo et al. (2006), Costa et al. (2008), and Padilha et al. (2016) in populations of Holstein cows reared in Brazil with an average of eight test-day records per cow. Therefore, the reduction of one or more test-day records caused reduction in the estimated genetic variance.

Variations in heritability estimates over lactation have been associated with different trends in permanent genetic and environmental variances (Cobuci et al., 2011). During lactation, while the permanent environmental variance decreased until the ninth month, the additive genetic variability gradually increased (Fig. 1), resulting in increasing heritability estimates after the first month of lactation. In the tenth month, the heritability was a slightly lower, which is due to the increase in the permanent environment variance in this period that is commonly the month when lactation ends.

Changes in the estimates of genetic parameters are common. Dionello et al. (2006) and Meseret et al. (2015) pointed to some reasons such as population size, breed, breeding conditions, selection systems, and the editing imposed on data used for genetic evaluation. In the present study, the greatest genetic variability was observed in P10 (Fig. 1) and was related to the largest number of herds and sires in this population (Table 1), suggesting that a decrease in the number of test-day records may impair genetic selection and decrease genetic gain.

The values and graphic formats of permanent environment correlation for all subpopulations are similar to that reported by Costa et al. (2008) for Holstein herds in Brazil and Torshizi et al. (2013) for animals of the Holstein breed raised in Iran, demonstrating that the reduction in the number of test-day records did not alter the permanent environment correlation between the months of records.

The greater genetic correlation between milk production on test-day on the different days throughout lactation in P10 may be related to the greater genetic variance estimated in this subpopulation (Fig. 1). There are generally no reports of such high genetic correlations between such distant periods (days). Usually, the most reported values and behaviors (Araújo et al., 2006; Meseret et al., 2015; Padilha et al., 2016) are similar to those reported in populations with reduced number of test-day records (such as occurred from P4 to P9).

**Reliability of estimated breeding value**

Comparing the average reliability of the breeding value for 305-day milk yield, for all situation considered in Table 2, the average reliability was higher in P10 and lower in P4, since as the number of test-day records reduced, the average reliability of the breeding value also decreased (between 22 and 29%). The greatest reduction occurred in the decrease of one record, which is related to the greater amount of phenotypic and genealogical information (Table 1). This result is particularly important for breeders, demonstrating that a
possible reduction in the number of records compromises the credibility of the genetic evaluation.

This reduction in reliability goes in the opposite direction to that sought by breeders, which is to increase the reliability of the genetic value, especially in young animals, reducing the generation interval that is achieved with the use of genomic information in genetic evaluations; the use of genomic information results in an increase of 8 to 43% in the reliability of the estimated breeding values (VanRaden et al., 2009).

This gain in reliability prompted research into the need for phenotypic records. Gonzalez-Recio et al. (2014) reported that a prolonged reduction in the availability of new phenotypes would have a negative effect on the reliability of improvement values in the future.

Thus, according to Gonzalez-Recio et al. (2014), it is necessary to ensure a continuous flow of phenotype collections to ensure predictive ability in the reference population, and that a financial compensation or payment for continued phenotype collection may be justifiable in view of the increased profitability provided by genetic improvement in the same population.

This demonstrates that if this practice is adopted, it will result in less reliable genetic values and a smaller number of animals, especially females with genetic evaluation. In the long term, this can compromise the results of a genetic improvement program. The same results are expected for other characteristics normally evaluated in dairy cattle.

However, other strategies for reducing costs may be applied, such as governmental financing, remote monitoring, and audit of milking control performed by farms without official controller from an association. Other ways to reduce the cost of milk control could be the adoption of subsidies and rise in the milk price for producers who carry out official milk control.

The alternating monthly milk test-day record (am/pm) is also a practice that can be used. Teixeira et al. (2006) reported a phenotypic correlation of 0.99 between observed and predicted milk yield and reported that potential advantages of this procedure compared with monthly test-day record would be reduction in the controller’s stay on the farm, less interference in the farm’s routine, lower cost to the breeder, and greater number of herds in control.

When considering sires with a similar number of daughters, the average reliability of the breeding value of these sires decreased as the number of test-day records was lower, with the greatest decrease in reliability observed with the decrease of one record, average of 16%, as shown in Fig. 2 (P10×P9).

The greater reliability of the estimate in the P10 population is due to the higher value of genetic variation in P10 (Fig. 1), because the calculation formula takes into account the value of genetic variation in the denominator (Misztal and Wiggans, 1988); therefore, a greater additive genetic variance results in greater reliability of the breeding value. Padilha et al. (2017) reported that when restrictions are applied in the genetic evaluation and there is a significant decrease in the number of sires and cows in the data sets, this can result in a decrease in the reliability of the estimated breeding values for 305-day milk yield.

The reliability of estimated breeding value is related to a larger population and availability of data on lactation (Oliveira et al., 2019a). Traits with low heritability value, that is, little genetic variation, also have low reliability of breeding values, requiring a number of animals or phenotypic measures to estimate the breeding value more reliably, which would not be necessary if the trait had high heritability (Oliveira et al., 2019b).

The reliability of the breeding value was superior for sires with the highest number of daughters, which was already reported by Santos et al. (2014). As this result was already expected, the bulls were divided into classes by number of daughters (Fig. 2) for a better discrimination of the differences and comparison between the evaluated subpopulations. This indicated a decrease in the number of sires in classes with a greater number of daughters (> 75) and with high reliability, which may result in a mistaken selection of the best sires. The small number of bulls with high reliability may become a problem in the future; if these are widely used, it may result in increased inbreeding, if it is not monitored in matings.

**Spearman’s ranking correlation**

Lower Spearman’s correlations (Fig. 3) were reported when comparing the selection of SB5 compared with SB20, and in DB2 compared with DC85, in all populations. This fact is possibly related to the selection differential used. That is particularly important given the fact that males have a greater number of offspring due to the use of artificial insemination, and thus spreading their genetic material more than females.

Sousa Júnior et al. (2014) working with weekly, monthly, bimonthly, and quarterly milk test-day records, mentioned that, although the coincidences are reasonable, the percentages indicate that there were divergences in the choice of the best sires and, consequently, in their classification.

The smallest Spearman’s correlations and coincidence from P10 with the others (P9 to P4) show that the decrease in one milk test-day record results in differences in the classification of the best animals, which can result in mistakes in the selection of the best animals for reproduction.
Some studies showed changes in the ranking of animals when different models were used to estimated breeding value. In Brazilian Holstein, when the best sires of 20%, 10%, 5%, and 2% were selected through different random regression models, the number of coincidences decreased (Bignardi et al., 2009), and for Guzerá cattle, when comparing the conventional model with random regression models, ranking correlations between 0.66 and 0.76 for the 5% best sires and 68 and 70% for the 20% best cows were reported (Santos et al., 2014). These changes demonstrate the importance of selecting a suitable model for genetic evaluation, and the same can be inferred for the data used in genetic evaluation.

In conclusion, the genetic evaluation performed under random regression models with the absence of one of 10 monthly test-day records per cow per lactation (incomplete lactation) is different from that in which all monthly test-day records are measured (complete lactation). This difference is observed sharply in terms of alteration in the ranking of animals and reduction in the reliability of estimated breeding values. This goes in the opposite direction to what is sought worldwide by the area of animal breeding, which seeks to use technologies such as insertion of genomic data to increase the reliability of the estimated breeding value. These unwanted impacts reinforce the need to maintain the monthly test-day milk record in Holstein cows, and also in other breeds raised in Brazil, regardless of the difficulties commonly encountered in carrying it out in Brazilian herds.

Author contribution GLF, FCB, PRNR, and JAC conceived and designed research. GLF, VTM, FCB, and MMO conducted analyzed data. GLF, FCB, PRNR, RN, and JAC wrote the manuscript. All authors read and approved the manuscript.

Funding This study was supported by the Coordenação de Aperfeiçoamento de Pessoal de Nível Superior (CAPES; Finance Code 001) – Brasil.

Data availability The datasets analyzed during the current study are not publicly available due to the data provider’s request.

Declarations

Ethics approval This is an observational study. The Federal University of Santa Maria—UFSM Research Ethics Committee has confirmed that no ethical approval is required.

Consent to participate Informed consent was obtained from all individual participants included in the study.

Conflict of interest The authors declare no competing interests.

References

Araújo, C.V., Torres, R.D.A., Costa, C.N., Torres Filho, R.D.A., Araújo, S.I., Lopes, P.S., Regazzi, A.J., Pereira, C.S. and Samiento, J.L.R., 2006. Uso de modelos de regressão aleatória para descrever a variação genética da produção de leite na raça Holandes da Revista Brasileira de Zootecnia, 35, 975–981
Biasussi, L.D.O., Cobucci, J.A., Costa, C.N., Roberto, P., Rotaro, N., Neto, J.B. and Cardoso, L.L., 2011. Genetic parameters for production traits in primiparous Holstein cows estimated by random regression models Revista Brasileira de Zootecnia, 40, 85–94
Bignardi, A.B., El Faro, L., Lucía, V., Fernando, P., Galvão, L. and Albuquerque, D., 2009. Random regression models to estimate test-day milk yield genetic parameters Holstein cows in South-eastern Brazil Livestock Science, 123, 1–7
Cardoso, V.L., Cassoli, D.L., Guilhermino, M.M., Machado, P.F., Nogueira, J.R. and Freitas, M.A.R., 2005. Análise econômica de esquemas alternativos de controle leiteiro Arquivo Brasileiro de Medicina Veterinária e Zootecnia, 57, 85–92
Cobucci, J.A., Costa, C.N., Braccini-Neto, J. and Freitas, A.F., 2011. Genetic parameters for milk production by using random regression models with different alternatives of fixed regression modeling Revista Brasileira de Zootecnia, 40, 557–567
Costa, C.N., Manoel, C., Melo, R. De, Packer, I.U., Freitas, A.F. De, Teixeira, N.M. and Cobucci, J.A., 2008. Genetic parameters for test day milk yield of first lactation Holstein cows estimated by random regression using Legendre polynomials Parâmetros genéticos para a produção de leite do dia do controle da primeira lactação de Brazilian Journal of Animal Science, 37, 602–608
Damane, M.M., Fozi, M.A. and Mehrgardi, A.A., 2016. Influence of milking frequency on genetic parameters associated with the milk production in the first and second lactations of Iranian Holstein dairy cows using random regression test day models Journal of Animal Science and Technology, 58, 5
Dionello, N.J.L., Silva, C.A., Costa, C.N. and Cobucci, J.A., 2006. Estimation of genetic parameters for test day milk yield of first lactation Jersey cows using repeatability and random regression models Revista Brasileira De Zootecnia-Brazilian Journal of Animal Science, 35, 1646–1652
Dzomba, E., Nephawe, K., Maiwashe, A., Cloete, S., Chimonyo, M., Banga, C., Muller, C. and Dzama, K., 2011. Random regression test-day model for the analysis of dairy cattle production data in South Africa: Creating the framework South African Journal of Animal Science, 40, 273–284
Gonzalez-Recio, O., Coffey, M.P. and Pryce, J.E., 2014. On the value of the phenotypes in the genomic era Journal of Dairy Science, 97, 7905–7915
Kirkpatrick, M., Hill, W.G. and Thompson, R., 1994. Estimating the covariance structure of traits during growth and aging, illustrated with lactations in dairy cattle Genetic research, 64, 57–69
Meseret, S. and Negussie, E., 2017. Genetic parameters for test-day milk yield in tropical Holstein Friesian cattle fitting a multiple-lactation random regression animal model South African Journal of Animal Science, 47, 352–361
Meseret, S., Tamir, B., Gebreyohannes, G., Lidauer, M. and Negus-sie, E., 2015. Genetic Analysis of Milk Yield in First-Lactation Holstein Friesian in Ethiopia: A Lactation Average vs Random Regression Test-Day Model Analysis Asian Australas. J. Anim. Sci., 28, 1226–1234
Misztal, I., Tsuruta, S., Strabel, T., Auveray, B., Druet, T. and Lee, D.H., 2002. Blup90 and Related Programs (Bgf90) In: 7th World Congress on Genetics Applied to Livestock Production, 2001–2002
Misztal, I. and Wiggans, G.R., 1988. Approximation of Prediction Error Variance in Large-Scale Animal Models Journal of Dairy Science, 71, 27–32
Mohammadi, A., Alijani, S. and Daghighkia, H., 2014. Comparison of different polynomial functions in Random Regression Model for milk production traits of Iranian Holstein dairy cattle Annals of Animal Science, 14, 55–68
Naserkheil, M., Seyed, R.M.A., Ardeshir, N.J., Son, J. and Lee, D., 2016. Random regression models using legendre polynomials to estimate genetic parameters for test-day milk protein yields in iranian holstein dairy cattle Asian-Australasian Journal of Animal Sciences, 29, 1682–1687
Ojango, J.M.K., Wasike, C.B., Enahoro, D.K. and Okeyo, A.M., 2016. Dairy production systems and the adoption of genetic and breeding technologies in Tanzania, Kenya, India and Nicaragua Animal Genetic Resources/Resources génétiques animales/Recursos genéticos animales, 59, 81–95
Oliveira, H.R., Brito, L., Silva, F., Lourenco, D., Jamrozik, J. and Schenkel, F., 2019a. Genomic prediction of lactation curves for milk, fat, protein, and somatic cell score in Holstein cattle Journal of Dairy Science, 102, 452–463
Oliveira, H.R., Lourenco, D.A.L., Masuda, Y., Misztal, I., Tsuruta, S., Jamrozik, J., Brito, L.F., Silva, F.F. and Schenkel, F.S., 2019b. Application of single-step genomic evaluation using multiple-trait random regression test-day models in dairy cattle. Journal of Dairy Science 1–13
Olori, V.E., Hill, W.G., McGuirk, B.J. and Brotherstone, S., 1999. Estimating variance components for test day milk records by restricted maximum likelihood with a random regression animal model Livestock Production Science, 61, 53–63
Padilha, A.H., Cobuci, J.A., Costa, C.N. and Neto, J.B., 2016. Random regression models are suitable to substitute the traditional 305-day lactation model in genetic evaluations of holstein cattle in Brazil Asian-Australasian Journal of Animal Sciences, 29, 759–767
Padilha, A.H., Costa, C.N., Braccini Neto, J., Daltro, D. dos S. and Cobuci, J.A., 2017. Selecting random regression models under different minimum number of test day records Livestock Science, 199, 69–73
Rural Pecuária, 2016. Rural Pecuária (São José do Rio Preto) Santos, D.J.A., Peixoto, M.G.C.D., Aspilcueta Borquis, R.R., Panetto, J.C.C., El Faro, L. and Tonhati, H., 2014. Predicting breeding values for milk yield of Guzerá (Bos indicus) cows using random regression models Livestock Science, 167, 41–50
Sousa Júnior, S.C. de, Faro, L. El, Bignardi, A.B., Cardoso, V.L., Machado, P.F. and Albuquerque, L.G. de, 2014. Aplicação de modelos de regressão aleatória utilizando diferentes estruturas de dados Ciência Rural, 44, 2058–2063
Teixeira, N.M., Freitas, A.F., Costa, C.N., Cobuci, J.A. and Barra, R.B., 2006. Controles alternados das ordenhas da manhã e da tarde para estimação da produção de leite na lactação Arquivo Brasileiro de Medicina Veterinária e Zootecnia, 58, 394–400
Torshizi, M.E., Aslamenejad, A.A., Nassiri, M.R., Farhangfar, H., Solkner, J., Kovac, M., Meszaros, G. and Malovrh, S., 2013. Analysis of Test Day Milk Yield by Random Regression Models and Evaluation of Persistency in Iranian Dairy Cows Iranian journal of applied animal science, 3, 67–76
VanRaden, P.M., Van Tassell, C.P., Wiggans, G.R., Sonstegard, T.S., Schnabel, R.D., Taylor, J.F. and Schenkel, F.S., 2009. Invited review: Reliability of genomic predictions for North American Holstein bulls Journal of Dairy Science, 92, 16–24
Zavadilová, L., Jamrozik, J. and Schaeffer, L.R., 2005. Genetic parameters for test-day model with random regressions for production traits of Czech Holstein cattle Czech Journal of Animal Science, 50, 142–154

Publisher’s note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Springer Nature or its licensor holds exclusive rights to this article under a publishing agreement with the author(s) or other rightsholder(s); author self-archiving of the accepted manuscript version of this article is solely governed by the terms of such publishing agreement and applicable law.