

## Genetic evaluation of Alpine goats using different milk control intervals

F.G. Silva, R.A. Torres, L.F. Brito, L.P. Silva, G.R.O. Menezes,  
L.C. Brito, R.F. Euclides and M.T. Rodrigues

Departamento de Zootecnia, Universidade Federal de Viçosa, Viçosa, MG, Brasil

Corresponding author: F.G. Silva  
E-mail: felipepheu@yahoo.com.br

Genet. Mol. Res. 12 (3): 2455-2464 (2013)  
Received August 13, 2012  
Accepted January 26, 2013  
Published July 24, 2013  
DOI <http://dx.doi.org/10.4238/2013.July.24.1>

**ABSTRACT.** The objective of this study was to compare the results of genetic evaluations by using different milk control intervals to reduce the cost of milk yield controls without harming the quality of genetic evaluation of the animals. We analyzed test day milk yield data from the Goat Sector of Universidade Federal de Viçosa. After editing and checking for errors in the database, there were 20,710 records of test day milk yield for the 667 first lactations of Alpine goats, constituting the complete file, with 7-day control intervals. Information on specific weeks was excluded from the complete file to create files with data on control intervals of 15, 21, and 28 days. The RENPED program was used to recode the pedigree and data files and correct pedigree errors; the WOMBAT program was used for genetic evaluations of the 4 files. The following comparison criteria of analysis results were used: logarithm of the function of the restricted maximum likelihood, length of the analyses in seconds, Pearson and Spearman correlations, and common elimination percentage among the areas below the regression curve of the genetic values of the animals. Overall, it is recommended that a 7-day interval among milk controls should be used in breeding programs and farms with a high technical level. Intervals of 14 and 21 days can achieve satisfactory

results combined with a lower data collection cost for farms with an average-to-low technical level, less effective size, and genetic variability that depend on external genetic material for genetic improvement.

**Key words:** Percentage of in common disposal; Legendre polynomials; Spearman correlation

## INTRODUCTION

The principal measured trait in dairy goats is test day milk yield (TDMY), which is the sum of the milking yield in 1 day. Data collection of TDMY in a short interval implies an increase in the qualified labor costs and personal transportation, whereas longer intervals contribute to a loss of quality in the genetic evaluations of the animals.

Cardoso et al. (2005) carried out a study specifying the costs of different designs of milk controls for cows in São Paulo State, Brazil. Such studies compare monthly, bimonthly, and trimonthly intervals, the cost of data collection and milk samples, analyses of the components, and transfer. Overall, they show that the associated costs can be reduced by adopting designs with longer intervals among the controls; in addition, bigger herds have a lower unitary cost and milk price is an important factor when choosing the design.

TDMY has been used for a number of decades to estimate the accumulated milk yield (AMY), which is evaluated by using many computer algorithms under a mixed model (Henderson, 1949). More recently, TDMY has been used in a more direct manner to select dairy goats by using the random regression models (RRMs) proposed by Henderson (1982). The utilization of RRM has advantages for modeling because they use information better, are less sensitive to data loss, consider variations of production within lactation, and work with more information, making it possible to obtain estimates of parameters and to predict more precise and accurate genetic values, resulting in greater genetic gains.

According to Brito et al. (2010), the use of milk controls is not very common among Brazilian goat farms due to their high cost. Other factors for their low use are the instability of the activity with a great range of milk demand and price, difficulty of farmers to use the generated data for animal selection, and the lack of practical results over the short and medium term, all of which discourage the majority of goat farmers adopting this practice.

Although the use of RRM for the genetic evaluation of TDMY provides safer estimates than analyses of AMY, it is difficult to choose the best animals based on the parameters of the generated curve. The use of genetic value on the greatest heritability point in the lactation curve as a selection criterion presents a reduction on the efficiency of the response correlated with complementary periods, according to the theory presented by Falconer (1981), due to the reduction of genetic correlation determined by linked genes. An interesting alternative to the common methods is the selection of animals by using the area below the regression of additive genetic deviations, which is specific to each animal within the appropriated interval to the production system of the evaluated herd.

The objective of this study was to compare the results of genetic evaluation by using different control intervals to identify an option that reduces costs without harming the quality of the genetic evaluation of the animals.

## MATERIAL AND METHODS

TDMY data from the Goat Sector of Universidade Federal de Viçosa (Viçosa, MG, Brazil) were used. The animals were raised in a free stall system and fed diets based on corn silage and hay as roughage and a concentrated mixture provided according to the nutritional needs of the animals.

Milk controls were collected using machine milking that was performed twice a day on a weekly basis. The morning collection started at 6:00 am and the afternoon collection started at 14:00 pm. TDMY is the sum of the milk collected in these 2 periods.

After editing and checking the database for errors, there were still 20,710 records of TDMY for 667 first lactations of Alpine goats, which constituted the complete file (i7), with 7-day intervals between the milk controls. Information on specific weeks was excluded from the complete file to create data files with intervals of 14 (i14), 21 (i21), and 28 (i28) days among milk controls. The RENPED program (Silva, 2011) was used to recode the pedigree and data files and to correct pedigree errors. The WOMBAT program (Meyer, 2006-2009) was used for genetic evaluations of the animals in the 4 files.

For data analyses of all files, the same single-trait animal model of random regression with orthogonal Legendre polynomials was used for fixed, additive genetic, and permanent environment regressions for orders 4, 2, and 7, respectively, in addition to 5 classes of residual variance. According to the following equation:

$$y_{ij} = EF_i + \sum_{m=0}^{4-1} b_m \varnothing_m(t_{ij}) + \sum_{m=0}^{2-1} \alpha_{im} \varnothing_m(t_{ij}) + \sum_{m=0}^{7-1} \gamma_{im} \varnothing_m(t_{ij}) + \varepsilon_{ij}$$

in which,  $y_{ij}$  = milk yield on the  $j^{\text{th}}$  control of goat  $i$ ;  $EF_i$  = set of fixed effects, which consisted of the genetic grouping (Alpine > 90% = 1; 90% ≥ Alpine > 80% = 2; 80% ≥ Alpine > 70% = 3; 70% ≥ Alpine > 60% = 4; and 60% ≥ Alpine > 50% = 5), season-year (rainy season from October to March and dry season in the other months, combined with 1998-2010), and type of kidding (0 to 3 kids), and age of goat at kidding was included as a co-variable with linear and quadratic effects;  $b_m$  = coefficient and  $m$  regression of TDMY on the Legendre polynomial to the model population average curve;  $\alpha_{im}$  and  $\gamma_{im}$  =  $m$  coefficients of the additive genetic regression and permanent environment, respectively, for goat  $i$ ; 4, 2, and 7 = orthogonal Legendre polynomial orders;  $t_{ij}$  is the control variable, lactation week of the  $i^{\text{th}}$  goat standardized for the interval from -1 to 1 according to Kirkpatrick et al. (1990);  $\varnothing_m(t_{ij})$  is the Legendre polynomial function for parameter  $m$  evaluated for age  $t_{ij}$ ; and  $\varepsilon_{ij}$  is the specific effects of each observation not explained by the correction factors or regression of the model.

The following criteria were compared among the controls: logarithm of the restricted maximum likelihood function and number of conditions for genetic regression (NCGR) and permanent environment (NCPE), length of analyses in seconds (T), Pearson and Spearman correlations, and the percentage of disposals in common (PDC) among the areas below the curve of genetic values of the animals.

The condition number is the division of the highest eigenvalue by the lowest eigenvalue associated with the parameters; high values indicate problems of multicollinearity. More

details can be seen in Montgomery and Peck (1981).

Disposal levels of 10, 20, 30, and 40% were defined to create the PDC, generating the 4 comparison criteria PDC<sub>10%</sub>, PDC<sub>20%</sub>, PDC<sub>30%</sub>, and PDC<sub>40%</sub>, respectively. The animals were evaluated genetically 4 times, once for each file (i7, i14, i21, and i28), by using the WOMBAT program (Meyer, 2006-2009). The area below the genetic curve of each animal, for the periods of 7 to 210 days and 7 to 290 days (Silva, 2011), was calculated from the exit file of WOMBAT (RnSoln\_animal.dat) by using the RENPED program, generating accumulated genetic values (a210 and a290, respectively). Animals with the lowest accumulated genetic value within the analyzed disposal level were separated and the percentage of animals in common among the files was then calculated.

## RESULTS AND DISCUSSION

Table 1 presents the number of records for TDMY, logarithm of the restricted maximum likelihood function (LogL), NCGR, NCPE, and T. All analyses were short, indicating that all of the files had a volume of data that could impair analyses using the F4A2EP7H5 model. The LogL results indicated that its increment occurs proportionally to the volume of the analyzed data and the utilization of intervals among the milk controls that are different from 7 will cause less accurate estimates of the genetic values of the animals.

**Table 1.** Number of records (NR), number of conditions for genetic regression (NCGR) and permanent environment (NCPE), logarithm of the restricted maximum likelihood function (LogL), and length of analysis in seconds (T) for analyses of files with intervals (i) of 7, 14, 21, and 28 days.

File	NR	NCGR	NCPE	LogL	T (s)
i7	20,670	12.23	109.60	5518.98	18.16
i14	10,168	12.57	263.32	1812.54	9.17
i21	6,839	20.32	449.12	536.58	8.41
i28	3,290	200.22	*	200.22	10.86

\*Non-estimable.

The NCGR results indicate strong multicollinearity for the additive genetic regression analysis of the i28 file and weak multicollinearity for the other analyses. NCPE from the analyses of the i28 file could not be performed because of the lack of an accurate estimate of the lower eigenvalue, which was indicated as zero. For the other analyses, multicollinearity was strong and tended to increase as the interval among the milk controls increased.

Table 2 presents the descriptive statistics for the accumulated genetic values a210 and a290 for files i7, i14, i21, and i28. The means, amplitudes, and standard deviations were greater for a290 in relation to a210, which may be due to the greater heritability of TDMY from 211 to 290 days or because of problems estimating the genetic values closer to end of the lactation period due to the lower amount of data during this period.

Table 3 presents the Pearson correlation for the accumulated genetic values from 7 to 210 days (above the main diagonal) and 7 to 290 days (below the main diagonal) among the 4 intervals of the controls analyzed. The Pearson correlation among the genetic values of the animals, obtained by analyses of the 4 files, indicates estimates of the genetic correlation of milk production among the different intervals of the milk controls. The results expressed in Table 3 indicate that the animals received similar genetic values from the analyses considering

the intervals among the milk controls from 14 to 21 days. The genetic values from the analyses considering 28-day intervals were further from the results of the other analyses.

**Table 2.** Number of animals (NA), minimal, mean, and maximum and standard deviation (SD) and coefficient of variation (CV), for accumulated genetic values (AGV) considering different intervals (i) among controls.

File	AGV	NA	Minimal	Mean	Maximum	SD	CV
i7	a210	905	-110.38	9.79	186.18	50.15	5.12
i7	a290	905	-153.66	13.62	259.20	69.82	5.12
i14	a210	905	-128.81	11.56	186.32	48.39	4.19
i14	a290	905	-179.32	16.09	259.39	67.37	4.19
i21	a210	905	-122.24	10.30	208.09	48.98	4.76
i21	a290	905	-170.18	14.34	289.69	68.19	4.76
i28	a210	905	-155.68	13.13	215.08	52.53	4.00
i28	a290	905	-216.73	18.28	299.42	73.13	4.00

**Table 3.** Pearson correlation for trait a210 (above the main diagonal) and a290 (below the main diagonal).

	i7	i14	i21	i28
i7	1	0.9454	0.9016	0.7662
i14	0.9428	1	0.9412	0.8270
i21	0.9043	0.9434	1	0.8362
i28	0.7703	0.8333	0.8359	1

i = interval in days.

Table 4 presents the Spearman correlation for the accumulated genetic values from 7 to 210 days (above the main diagonal) and from 7 to 290 days (below the main diagonal) among the 4 intervals of the controls analyzed. The Spearman correlation among the genetic values of the animals, obtained by analyses of the 4 files, indicates the degree of correspondence among the animal classifications in distinct analyses. This measurement is very important when the classification of one animal indicates not only if this animal will be selected for reproduction but also at what intensity it will be used.

**Table 4.** Spearman correlation for trait a210 (above the main diagonal) and a290 (below the main diagonal).

	i7	i14	i21	i28
i7	1	0.9317	0.8943	0.7467
i14	0.9317	1	0.9374	0.8120
i21	0.8943	0.9374	1	0.8230
i28	0.7465	0.8119	0.8229	1

i = interval in days.

Tables 5, 6, 7, and 8 present the PDCs considering 10, 20, 30, and 40% of eliminated females per year, respectively. The results for the genetic accumulated value of 7 to 210 days and 7 to 290 days are expressed above and below the main diagonal, respectively. For both the accumulated genetic values, Tables 5, 6, 7, and 8 indicate that the use of any, with the exception of the 7-day interval, would harm the choice of the best animals. However, in low to intermediate technical systems, intervals of 14 to 21 days between the milk controls can be used to reduce production control costs.

**Table 5.** Percentage of in common disposal, considering 10% of females eliminated per year, for a210 (above the main diagonal) and a290 (below the main diagonal) among the 4 intervals analyzed.

	i7	i14	i21	i28
i7	-	82.61	69.57	51.09
i14	80.43	-	73.91	56.52
i21	68.48	75.00	-	64.13
i28	50.00	56.52	65.22	-

**Table 6.** Percentage of in common disposal, considering 20% of females eliminated per year, for a210 (above the main diagonal) and a290 (below the main diagonal) among the 4 intervals analyzed.

	i7	i14	i21	i28
i7	-	81.52	79.35	71.74
i14	79.67	-	90.22	75.00
i21	71.98	83.52	-	73.91
i28	60.99	65.38	70.33	-

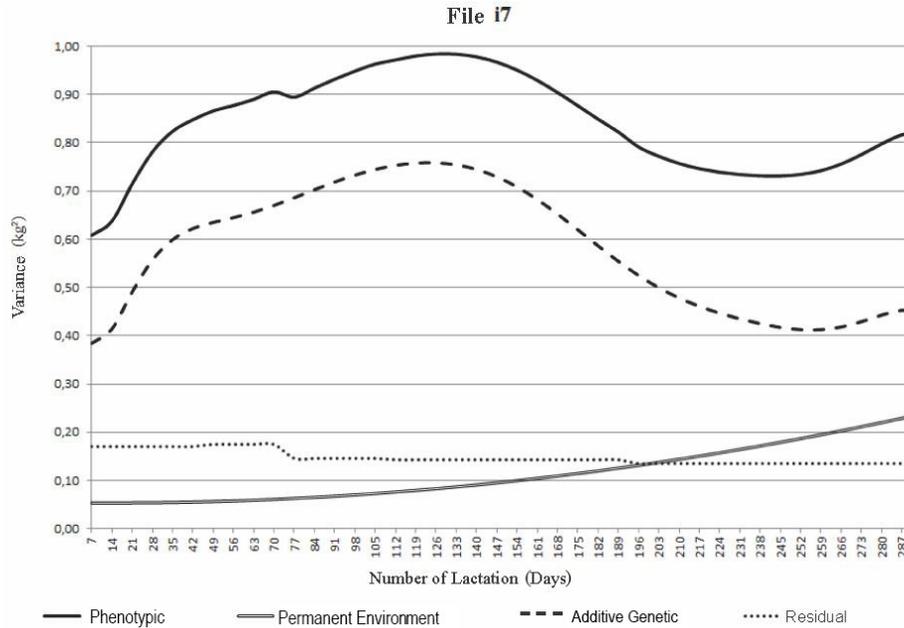
**Table 7.** Percentage of in common disposal, considering 30% of females eliminated per year, for a210 (above the main diagonal) and a290 (below the main diagonal) among the 4 intervals analyzed.

	i7	i14	i21	i28
i7	-	86.55	81.82	64.00
i14	86.55	-	87.27	67.64
i21	81.82	87.27	-	70.55
i28	64.00	67.64	70.55	-

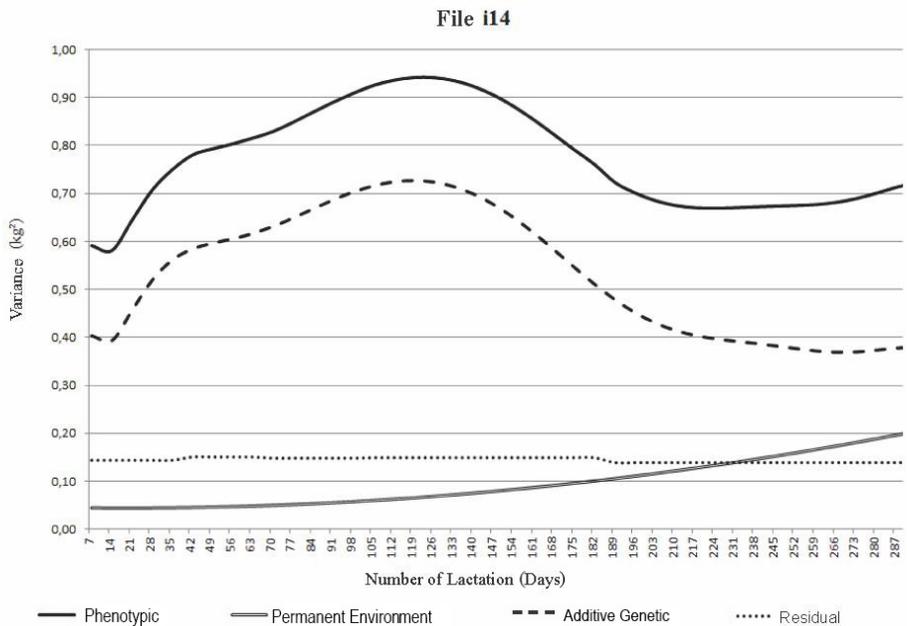
**Table 8.** Percentage of in common disposal, considering 40% of females eliminated per year, for a210 (above the main diagonal) and a290 (below the main diagonal) among the 4 intervals analyzed.

	i7	i14	i21	i28
i7	-	86.34	82.51	69.13
i14	86.34	-	86.89	70.77
i21	82.51	86.89	-	74.32
i28	69.13	70.77	74.32	-

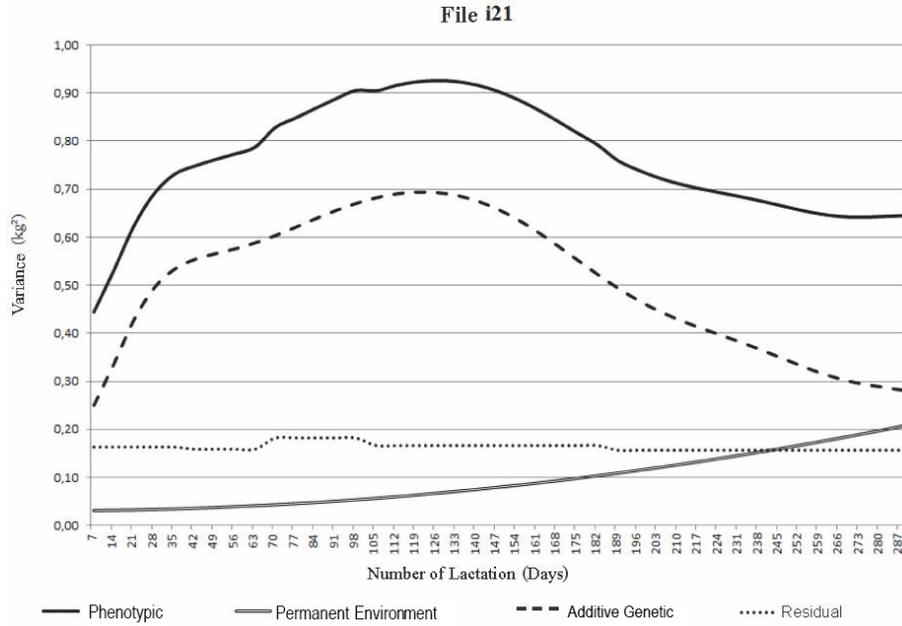
Figures 1, 2, 3, and 4 present the phenotypic, additive genetic, permanent environment, and residual variances of TDMY in relation to the number of days in lactation from the 7th to the 290th day, considering the different intervals of the controls evaluated. The results indicate the relatively similar behavior among the parameter estimates of TDMY for i14 and i21 in relation to i7. The greatest divergence occurred for the permanent environment variance of i21, which was close to 290 days. The estimates for i28 were very distinct from the estimates for i7, i14, and i21, indicating that long control intervals could harm the parameter estimates of TDMY.



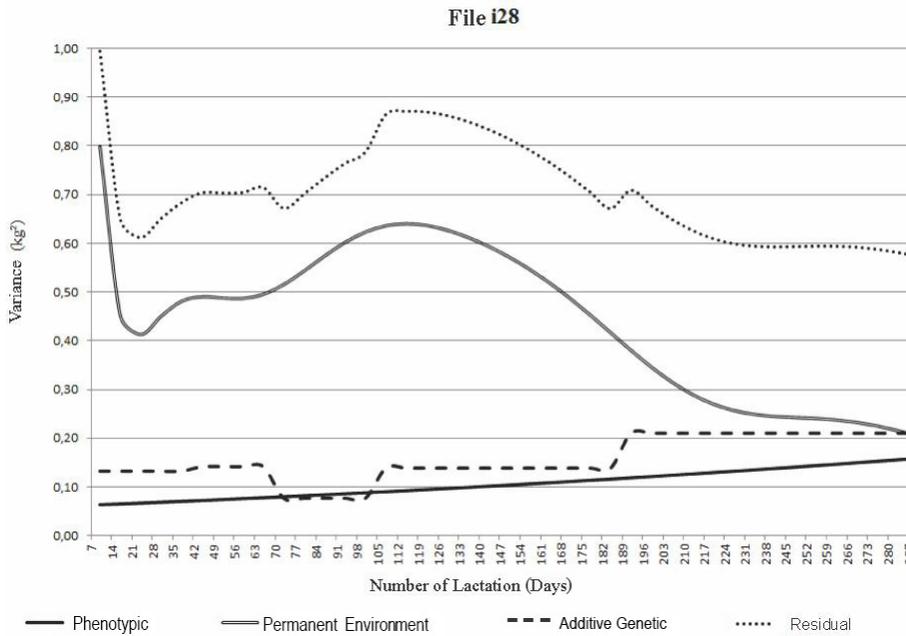
**Figure 1.** Estimates of phenotypic, additive genetic, permanent environment, and residual variances for test day milk yield using control intervals of 7 days, represented by files i7.



**Figure 2.** Estimates of phenotypic, additive genetic, permanent environment, and residual variances for test day milk yield using control intervals of 14 days, represented by files i14.

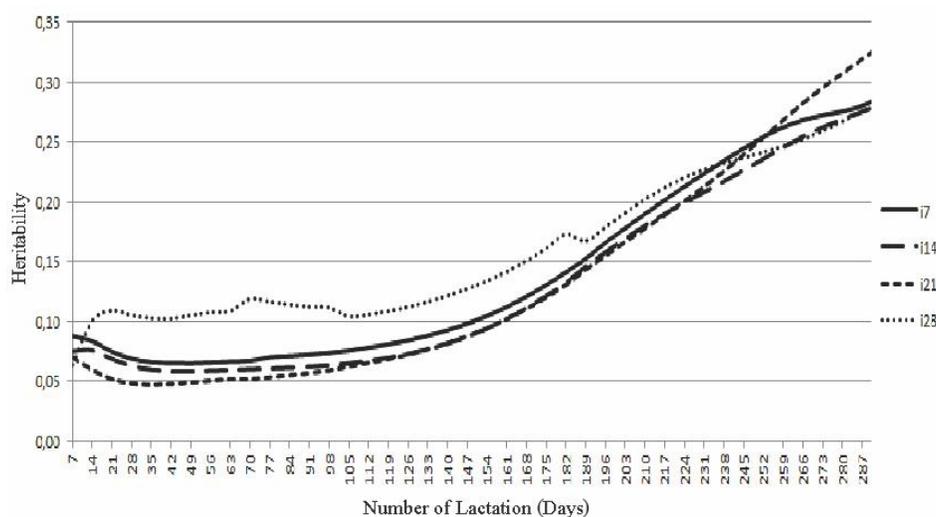


**Figure 3.** Estimates of phenotypic, additive genetic, permanent environment, and residual variances for test day milk yield using control intervals of 21 days, represented by files i21.



**Figure 4.** Estimates of phenotypic, additive genetic, permanent environment, and residual variances for test day milk yield using control intervals of 28 days, represented by files i28.

Figure 5 shows the heritability of TDMY as a function of the number of days in lactation from 7 to 290 days for the different intervals among the controls studied. The results show the similar behavior of heritability estimated for i14 and i21 in comparison to i7 until 210 days; after that time, the behavior of i21 changes. The behavior of heritability for i28 was very distinct from the behavior of i7, indicating that their estimates are unstable; this may be linked to the low precision and accuracy caused not only by the lower quantity of data but also by the longer interval among the measurements, which generally reduces the quality of the prediction for regression coefficients of the permanent environment curves, which are specific to each animal.



**Figure 5.** Heritability of test day milk yield in function of the number of days in lactation considering different intervals among milk controls.

## CONCLUSIONS

The results of the present study indicate that for breeding programs and farms with a high technical level, intervals of 7 days among milk controls are necessary for more precise and accurate analyses. For less technical farms, the results for milk control intervals of 14 or 21 days indicated that they may be a good alternative for reducing the cost of data collection.

## REFERENCES

- Brito LF, Caetano GC, Pereira BM, Silva FG, et al. (2010). Avaliação Genética de Cabras Leiteiras Usando Intervalos Alternativos de Controle Leiteiro e Produções Parciais. Anais do Simpósio de Integração Acadêmica, Universidade Federal de Viçosa, Viçosa.
- Cardoso VL, Cassoli LD, Guilhermino MM, Machado PF, et al. (2005). Análise econômica de esquemas alternativos de controle leiteiro. *Arq. Bras. Med. Vet. Zootec.* 57: 85-92.
- Falconer DS (1981). Introdução à Genética Quantitativa. Universidade Federal de Viçosa, Viçosa.
- Henderson CR (1949). Estimation of changes in herd environment. *J. Dairy Sci.* 32: 706.
- Henderson Jr CR (1982). Analysis of covariance in the mixed model: higher-level, nonhomogeneous, and random

- regressions. *Biometrics* 38: 623-640.
- Kirkpatrick M, Lofsvold D and Bulmer M (1990). Analysis of the inheritance, selection and evolution of growth trajectories. *Genetics* 124: 979-993.
- Meyer K (2006-2009). WOMBAT - A program for Mixed Model Analyses by Restricted Maximum Likelihood. Manual Version 1.0.
- Montgomery DC and Peck EA (1981). Introduction to Linear Regression Analysis. John Wiley & Sons, New York.
- Silva FG (2011). Estudo da Curva de Produção de Leite de Caprinos Utilizando Modelos de Regressão Aleatória. Master's thesis, Universidade Federal de Viçosa, Viçosa.