

## Estimation of heritability and genetic trend in populations at a physiological limit

T. R. Famula<sup>1</sup> and L. D. Van Vleck<sup>2</sup>

<sup>1</sup> Department of Animal Science, University of California, Davis, CA 95616, USA

<sup>2</sup> Agricultural Research Service, Roman L. Hruska Meat Animal Research Center, A218 Animal Sciences, University of Nebraska, Lincoln, NE 68583-0908, USA

Received October 10, 1989; Accepted January 22, 1990

Communicated by E. J. Eisen

**Summary.** Limits on physiological processes, though perhaps unknown, must exist. The reported simulations evaluate the effect of a physiological limit on the estimation of genetic parameters and genetic progress. Simulation experiments reveal no change in the estimate of heritability, even for limits as restrictive as 1.5 phenotypic standard deviations above the population mean. However, estimates of additive genetic and environmental variance shrink as limits on performance increase in severity. Simulated physiological limits do not affect the rate of genetic progress. However, absolute measures of estimated genetic change are less than those predicted by response equations.

**Key words:** Genetic trend – Heritability – Selection limit

### Introduction

Intuitively, every measure of animal production has a limit of performance. Most species of domestic livestock have not, however, reached such limits. Nevertheless, such limits to performance must exist. Evidence for the existence of limits is available in a few, isolated cases.

A recent estimate of genetic trend in racing performance of thoroughbred horses shows no change in record winning times over the past several decades (Gaffney and Cunningham 1988). Their results demonstrate a significant positive genetic trend, accompanied by retention of considerable additive genetic variance. Yet record times have failed to show a correlated change. Cunningham (1989) suggests that winners of classic horse races are close to a physiological limit in performance, and that a mechanism causing this limit is accumulation of blood lactic acid (Fregin and Thomas 1982).

Although debate may continue as to whether there is such a plausible mechanism, few would question that a limit to racing performance exists. Also, one may continue to debate the precise level of the limit. However, the existence of a limit is axiomatic.

Racing performance represents a rapid burst of physiological events. Yet we also assume that limits exist for extended physiological processes such as lactation or growth. Mathematical models of lactation and growth rely on Michaelis-Menten kinetics. Such models have obvious asymptotic limits (i.e.,  $V_{max}$ ). In lactation, most models suggest that mammary capacity for the uptake and utilization of metabolites is in excess of that expressed due to limitations in nutrient availability (Baldwin and Smith 1983). Once again, although knowledge of a precise limiting process is not available, present data suggest that asymptotic values do exist.

Statistically, a physiological limit represents a point beyond which observations have a zero probability of occurring. This phenomenon is unlike any other censoring process considered by animal breeders. Typical censoring involves the existence of a point beyond which observations can occur (with nonzero probability) but are not recorded.

Our purpose here is not to estimate such limits. Instead, we wish to study the impact of limits on estimates of future genetic progress. Bearing in mind the Gaffney and Cunningham (1988) data, what is the impact of a physiological limit on estimates of heritability and genetic trend? Is there substantial bias in estimates of genetic variance that will influence the interpretation of the data?

A natural assumption is to expect a decrease in estimates of variation, both genetic and phenotypic. Moreover, it is also natural to expect a decrease in estimates of genetic trend. However, a preliminary simulation study

(L. D. Van Vleck, personal communication 1985) showed that performance limits above the mean had no effect on regression estimates of heritability. Such unanticipated conclusions require an explanation. Hence, the objective of this report is to examine the impact of physiological limits on estimates of genetic variance and genetic trend.

## Materials and methods

The goal of these simulation experiments is to examine the joint impact of selection and physiological limits on estimates of genetic variance and genetic trend. Accordingly, the simulation involves two classes of data. The first class involves selection and the second serves as a randomly mated control. First, we consider the structure of the populations under selection.

### Population structure

We begin with the creation of 200 data sets (replicates) in each of three populations (total of 600 data sets). We define a population by the underlying true genetic and environmental variances, corresponding to heritabilities of 0.15, 0.25, and 0.35. Each data set contains 48 individuals. This is divided into three generations of 16 individuals per generation.

The base generation consists of 16 unrelated individuals with breeding values and phenotypes created from pseudo-normal random deviates. The phenotypic variance for each population is 100, with a mean of zero. The genetic and environmental variances depend on the underlying true heritability. With individual phenotype as the selection criterion, animals with the top 12 phenotypes are chosen as parents of the next generation. Four of the 12 are randomly assigned as males, with the remaining 8 as females. Each male receives two mates (chosen randomly) with two progeny created per mating to establish a second generation of 16 individuals. The breeding value of each offspring is the sum of a randomly generated sampling term (with a variance of one-half the additive genetic variance) and one-half the breeding value of each parent. To this breeding value a randomly generated environmental term (with mean zero and variance established by the underlying true heritability) is added to create each phenotype. The third generation of 16 individuals is created in a similar fashion.

The allocation of mates in generations one and two is random, except that mating of relatives is not permitted.

In populations undergoing selection, the choice of parentage is based on the individual's phenotype alone. Of course, in the populations not experiencing selection, the 12 parents of each generation are chosen strictly at random. Mates are assigned at random, with the exception that mates must not be related.

### Simulation of a physiological limit

The simplest way to simulate a physiological limit is to censor phenotypes above a specific value. Alternative, more complex models, would permit a physiological limit to vary from one individual to the next. The simulation involves five physiological limits. Each of the 200 simulated data sets is subjected to the five limits. Measured in phenotypic standard deviations, the limits imposed are 1.5  $\sigma$ , 1.75  $\sigma$ , 2.0  $\sigma$ , 2.25  $\sigma$ , and no limit. Because  $\sigma = 10$  for all the parameter sets, the actual limits are 15, 17.5, 20, and 22.5.

Once the 16 phenotypes are created within a generation, phenotypes above the imposed limit are censored to the value of the limit. For example, a phenotype of 16.8, in a population with a physiological limit of 15, is changed to 15. Censoring is done before the selection of parents and mating.

### Estimates of variances and genetic trends

A model for the phenotypes of the 48 individuals is

$$Y = Z a + e \quad (1)$$

where  $Y$  is the vector of 48 limited phenotypes,  $Z$  is a known incidence matrix ( $Z = I$ ),  $a$  is a random vector of 48 breeding values, and  $e$  is a vector of random residuals. Moreover,  $E[a] = E[e] = \mathbf{0}$  and

$$\text{Var} \begin{bmatrix} a \\ e \end{bmatrix} = \begin{bmatrix} A \sigma_A^2 & \mathbf{0} \\ \mathbf{0} & I \sigma_E^2 \end{bmatrix}$$

where  $A$  is a matrix of numerator relationships among the 48 individuals,  $\sigma_A^2$  is the additive genetic variance and  $\sigma_E^2$  is the residual variance. Note that  $a$  can be partitioned into three subvectors, each of order 16, as  $a' = [a'_1 | a'_2 | a'_3]$  where the vector  $a_i$  represents the breeding values of individuals of the  $i^{\text{th}}$  generation ( $i = 1, 2, 3$ ). Breeding values are predicted from the solution of

$$[Z'Z + A^{-1}k] \hat{a} = Z'Y \quad (2)$$

for  $k = \sigma_E^2 / \sigma_A^2$ .

Estimation of genetic trend is derived from computing the mean estimated breeding value of each generation:  $\mathbf{1}'\hat{a}_i/16$  for the  $i^{\text{th}}$  generation with  $\mathbf{1}$  a unit vector of order 16.

The second objective of this experiment was to assess the impact of physiological limits on estimates of heritability. Accordingly, we estimated  $\sigma_A^2$  and  $\sigma_E^2$  for each of the 200 simulated data sets. The method used was restricted maximum likelihood (REML), which involves the following quadratic forms,

$$\hat{a}' A^{-1} \hat{a}$$

and

$$\hat{e}' \hat{e} \quad (3)$$

for  $\hat{e} = y - Z\hat{a}$ . A more thorough discussion of REML and the algorithm used to solve for  $\sigma_A^2$  and  $\sigma_E^2$  can be found in Henderson (1984). For the iterative algorithm, we defined convergence as a change of less than 0.001 in both variance components between two successive rounds of iteration. The true values of the variance components were used as the starting values in the mixed model equations.

The imposition of a physiological limit to performance has one additional expected outcome. As should be clear from model (1), genotype and environment are independent. Yet the limit to performance should have the indirect effect of creating a negative genotype-by-environment covariance. Simply put, individuals with a breeding value well above average may have phenotypes above the physiological limit. When the data are forced to conform to the limit we are, in effect, adding a below-average environmental deviation.

To estimate this covariance term we chose a simple, approximate strategy; i.e., approximate in that the mixed model equations of line (2) do not incorporate a genotype-by-environment covariance. As such, estimates of this parameter cannot be made under a modified form of the algorithm discussed in Henderson (1984). Because this is a simulation, we could compute a second set of residuals using the true additive genetic values instead of the predicted genetic values. One can then compute the simple covariance between the true breeding value and these alternatively estimated residuals. Although approximate, this statistic does provide a simple means to detect the 'creation' of a genotype-by-environment covariance through the imposition of a limit to performance.

**Table 1.** Mean estimates of genetic and residual variances in populations with physiological limits and no selection

Estimates	Limit					
	True value	None	2.25 SD	2.00 SD	1.75 SD	1.50 SD
True heritability =0.35						
Additive variance	35	35.95	35.56	34.75	33.66	32.05
Residual variance	65	63.91	63.23	61.77	59.85	56.98
Total variance	100	99.86	98.79	96.52	93.51	89.03
Heritability	0.35	0.36	0.36	0.36	0.36	0.36
True heritability =0.25						
Additive variance	25	24.15	23.71	23.01	21.98	21.47
Residual variance	75	76.49	75.07	72.86	69.59	67.99
Total variance	100	100.64	98.78	95.87	91.57	89.46
Heritability	0.25	0.24	0.24	0.24	0.24	0.24
True heritability =0.15						
Additive variance	15	15.85	15.67	15.34	13.97	13.38
Residual variance	85	83.21	82.29	80.52	79.19	75.80
Total variance	100	99.06	97.96	95.86	93.16	89.18
Heritability	0.15	0.16	0.16	0.16	0.15	0.15

**Table 2.** Mean estimates of genetic and residual variances in populations with physiological limits and selection

Estimates	Limit					
	True value	None	2.25 SD	2.00 SD	1.75 SD	1.50 SD
True heritability =0.35						
Additive variance	35	35.85	35.38	33.05	31.70	30.15
Residual variance	65	63.73	62.90	61.37	58.86	55.98
Total variance	100	99.58	98.28	94.42	90.56	86.13
Heritability	0.35	0.36	0.36	0.35	0.35	0.35
True heritability =0.25						
Additive variance	25	25.77	24.17	23.21	21.84	20.17
Residual variance	75	73.36	72.50	69.62	65.52	63.86
Total variance	100	99.13	96.67	92.82	87.36	84.03
Heritability	0.25	0.26	0.25	0.25	0.25	0.24
True heritability =0.15						
Additive variance	15	15.98	14.55	14.10	13.31	11.72
Residual variance	85	83.91	82.44	79.87	75.45	71.96
Total variance	100	99.89	96.99	93.97	88.76	83.68
Heritability	0.15	0.16	0.15	0.15	0.15	0.14

## Results

Tables 1 and 2 present means of estimates of the additive genetic and residual variances for the three simulation parameter sets at each physiological limit. The means in Table 1 are from populations not experiencing selection. Table 2 presents similar statistics in populations undergoing selection on the simulated, and limited, trait.

A simple two-way analysis of variance (data not presented) reveals significant differences between data sets and limits for each of the variance components for all three parameter sets. However, estimates of heritability are not significantly different across limit classes for all

three simulation parameter sets. Moreover, the results are consistent, regardless of whether or not selection is practiced.

Upon closer examination of Table 1, we note a gradual decline in all variance estimates as the limit becomes more severe. Such a result is anticipated. Yet the stability of the heritability estimate is not so easily predicted.

Interestingly, in comparing estimates in Tables 1 and 2, the decline in variance estimates across limits is greater in populations experiencing selection. The consensus of simulation and analytic results suggests that REML estimates of variance components are not biased by specific forms of selection when 'tied' to an unselected base pop-

**Table 3.** Mean estimates of breeding value in populations with physiological limits and no selection (including the percentage of individuals exceeding the limit)

Generation	Limit					
	True value	None	2.25 SD	2.00 SD	1.75 SD	1.50 SD
True heritability = 0.35						
1	0.02	0.03 (0)	0.02 (1.2)	0.00 (2.2)	-0.04 (4.3)	-0.11 (7.8)
2	-0.06	-0.07 (0)	-0.05 (0.9)	0.03 (2.2)	-0.09 (4.0)	-0.19 (6.4)
3	0.04	0.05 (0)	0.04 (1.3)	0.01 (2.3)	0.00 (4.0)	-0.09 (6.7)
True heritability = 0.25						
1	-0.01	0.00 (0)	-0.01 (1.3)	-0.02 (2.3)	0.06 (3.9)	-0.11 (6.8)
2	0.05	0.05 (0)	0.05 (1.3)	0.02 (2.2)	-0.03 (4.3)	-0.11 (7.0)
3	-0.05	-0.06 (0)	-0.05 (1.2)	-0.08 (2.3)	-0.13 (3.8)	-0.23 (6.6)
True heritability = 0.15						
1	0.00	-0.01 (0)	-0.02 (1.0)	-0.03 (2.0)	-0.05 (3.7)	-0.09 (6.9)
2	0.04	0.05 (0)	0.04 (1.2)	0.02 (2.2)	-0.01 (3.5)	-0.07 (6.3)
3	-0.03	-0.01 (0)	-0.02 (1.0)	-0.04 (2.1)	-0.08 (4.0)	-0.14 (6.2)

**Table 4.** Mean estimates of breeding value in populations with physiological limits and selection (including the percentage of individuals exceeding the limit)

Generation	Limit					
	True value	None	2.25 SD	2.00 SD	1.75 SD	1.50 SD
True heritability = 0.35						
1	0.01	0.05 (0)	0.03 (6.8)	0.01 (1.8)	-0.04 (3.5)	-0.14 (6.1)
2	1.41	1.44 (0)	1.39 (2.0)	1.35 (3.6)	1.27 (5.9)	1.14 (9.4)
3	2.81	2.78 (0)	2.73 (1.9)	2.67 (3.9)	2.56 (6.9)	2.39 (10.9)
True heritability = 0.25						
1	0.00	0.02 (0)	-0.01 (1.4)	-0.03 (2.5)	-0.07 (4.1)	-0.14 (7.0)
2	1.08	1.1 (0)	1.05 (1.6)	1.01 (3.0)	0.94 (5.4)	0.83 (8.8)
3	2.07	2.11 (0)	2.05 (2.4)	2.00 (4.2)	1.91 (7.1)	1.76 (10.9)
True heritability = 0.15						
1	0.01	0.02 (0)	0.01 (1.1)	0.00 (2.5)	-0.03 (4.1)	-0.07 (6.6)
2	0.66	0.66 (0)	0.64 (1.4)	0.62 (2.5)	0.57 (4.5)	0.50 (7.6)
3	1.14	1.18 (0)	1.16 (1.4)	1.13 (3.0)	1.07 (4.7)	0.98 (7.6)

ulation. Gianola et al. (1989) lend support to this argument by demonstrating that REML equations taking selection into account are identical to those ignoring selection. This is true, of course, only if the data are distributed normally. In those data sets with imposed limits, an assumption of normality is not justified. Hence, variance estimates in populations undergoing selection are not expected to be equivalent to estimates in randomly mated populations. The fact that estimates of heritability remain unaltered, however, is unanticipated.

As might be expected for populations not undergoing selection, estimates of phenotypic variance are similar for the three parameter sets, because all data are simulated with a phenotypic variance of 100, regardless of underlying heritability. Although there is a significant decline

with more severe physiological limits, the decline in phenotypic variance is similar for each parameter set regardless of heritability.

Tables 3 and 4 present the means of estimates of breeding values by generation for selected and unselected populations, respectively. Breeding value estimates are from the first round solutions of the mixed model equations [Eq. (2)]. In populations not undergoing selection, genetic change should be nonexistent, as is evident in the trend of the mean true breeding values. This result is repeated in the estimates from populations without a physiological limit. However, with the imposition of a performance limit, and as that limit becomes more severe, the means of estimates of breeding value decline. The change from one generation to the next, across lim-

**Table 5.** Mean estimates of genotype-by-environment covariance in populations with physiological limits and no selection

	Limit				
	None	2.25 SD	2.00 SD	1.75 SD	1.50 SD
True heritability 0.35	-0.02	-0.27	-0.57	-1.12	-1.98
True heritability 0.25	0.03	-0.27	-0.55	-0.98	-1.62
True heritability 0.15	-0.04	-0.16	-0.30	-0.50	-0.81

**Table 6.** Mean estimates of genotype-by-environment covariance in populations with physiological limits and selection

	Limit				
	None	2.25 SD	2.00 SD	1.75 SD	1.50 SD
True heritability 0.35	-0.04	-0.44	-0.83	-1.49	-2.46
True heritability 0.25	0.03	-0.45	-0.81	-1.34	-2.14
True heritability 0.15	0.03	-0.13	-0.30	-0.57	-0.97

its, is consistent. Only at the most constraining of limits is the genetic change per generation influenced. Yet even these changes are minimal.

Table 4 presents the means of estimates of breeding value in populations experiencing selection. The first column, representing the change in true breeding values of the simulated genotypes, accurately reflects the expected change in each population. Using order statistics, the mean of the top 12 of 16 standard normal variables is 0.40 (Beyer 1966). Accordingly, the expected genetic change is 1.40, 1.00, and 0.60 in those populations with heritabilities of 0.35, 0.25, and 0.15, respectively.

As in those populations not undergoing selection, the imposition of a performance limit forces a decrease in the estimated breeding values. However, the decline is no more severe in populations with selection than those without selection. Perhaps the most interesting observation is that, although a physiological limit brings about a decline in the mean of estimated breeding values, the differences between generations are unaffected. Thus, the expected differences between generations as a result of selection (the values of 1.40, 1.00, or 0.60) are observed regardless of the imposed performance limit. As with heritability estimates, the existence of a physiological limit does not influence the estimation of genetic trend. Of course, this conclusion is confined to short-term selection response in populations with moderate physiological limits. The experiments here do not permit extrapolation to more severe limits or measures of response beyond three generations.

Tables 5 and 6 present mean estimates of the genotype-by-environment covariance. As outlined earlier, before this experiment, imposition of physiological limits would have been expected to create a negative genotype-

by-environment covariance. As the tables illustrate, this is clearly the case. As the limits imposed become more restrictive, this covariance moves farther from zero. Of course, with no limit, the value of this covariance is approximately zero, whether there is selection or not. However, the mean absolute value of this covariance is greater in populations with selection than in those where parents are chosen at random. Whether these differences are statistically significant has not been tested.

## Discussion

Intuitively, the existence of an upper limit to performance would be expected to bias estimates of genetic parameters. Indeed, the results in Tables 1 and 2 show a significant decrease in estimates of genetic, environmental, and phenotypic variance (two-way analysis of variance not presented). Yet the decrease in genetic and environmental variances is such that the estimate of heritability is unchanged.

We see a similar result in the estimates of genetic trend. The imposition of a physiological limit, as noted above, reduces the range of phenotypes. As a result, predictions of large breeding value also decline. The effect of this reduction is to decrease the absolute estimates of genetic progress. For example, in populations with a true heritability of 0.25, Table 4 presents the third-generation mean breeding value as 1.76 when the imposed limit is 1.5 phenotypic standard deviations above the mean. Yet the true genetic mean is 2.07. The rate of estimated genetic change remains unaffected. The changes in mean breeding value across generations are consistent with changes in limit. Thus, although the ab-

solute magnitude of genetic change declines, the rate of genetic change is unaffected by physiological limits to performance.

No doubt, a critical value in generating these results is the number of animals that exceed the limit. Tables 3 and 4 present the percentage of individuals that were beyond the physiological limit and had censored phenotypes. Although substantial, these percentages are not large enough to distort the distribution of phenotypes. Thus, if only a small portion of individuals is affected by the limit, the impact on estimation of variances and genetic trends is less pronounced.

If this model of physiological limits is appropriate, the results of Gaffney and Cunningham (1988) for heritability and rates of genetic change should be unaffected by a potential physiological limit to racing performance. Of course, estimates of variance components and mean breeding values will be reduced by a limit to performance. This enforces the obvious notion that physiological limit does not limit the population mean. The impact of the limit is to reduce the performance of individuals at the extreme. The mean performance still has the capacity for genetic change. The rate of such genetic change is little affected by the limiting of extreme phenotypes.

One explanation for the continuation of genetic progress in traits constrained by physiological limits is the inability of the limit to change those individuals selected as parents. Although predictions of breeding value are reduced, the same individuals are chosen as parents, regardless of the limit. Recall from the design of the experiment that the same series of pseudo-random normal deviates was chosen for each limit class. Hence, the same data were simulated (prior to the imposition of a limit) in all cases. Note that the same mean of true breeding values was computed in each limit class. This is evidence that the same 12 individuals were chosen as parents in each generation of the 200 data sets. So, within the limits imposed, the impact on the actual breeding program was negligible over the three generations simulated.

One outcome of these experiments that has no simple explanation is the failure of REML to behave the same

on selected data as on unselected data. A comparison of means in Tables 1 and 2 reveals that REML estimates of variance components are not only influenced by limits to performance, but that this influence is exaggerated by selected data. Most likely, this difference in mean estimates is a function of the number of animals that are censored by the physiological limit. More animals are censored in populations undergoing selection than in randomly mated populations. This fact may be the cause of the decline in variance estimates.

The results presented in Table 1 suggest that REML cannot overcome the bias induced by censoring extreme phenotypes. Such an outcome is natural to assume from the outset of the experiment. Yet from the outset one might also assume that selection would not additionally bias estimates of variances. That selection, in this setting, can bias estimates of variance components via REML is, perhaps, the most interesting result of these simulations.

## References

- Baldwin RL, Smith NE (1983) Adaptation of metabolism to various conditions: milk production. In: Riis PM (ed) *Dynamic biochemistry of animal production*. Elsevier, Amsterdam, pp 359–388
- Beyer WH (ed) (1966) *Handbook of tables for probability and statistics*. The Chemical Rubber Co., Cleveland, p 115
- Cunningham EP (1989) Thoroughbred breeding. *Nature* 337:414–415
- Fregin GF, Thomas DP (1982) Cardiovascular response to exercise in the horse: a review. In: Snow DH, Persson SGB, Rose RJ (eds) *Equine exercise physiology*. Burlington Press, Cambridge, pp 76–90
- Gaffney B, Cunningham EP (1988) Estimation of genetic trend in racing performance of thoroughbred horses. *Nature* 332:722–724
- Gianola D, Fernando RL, Im S, Foulley JL (1989) Likelihood estimation of quantitative genetic parameters when selection occurs: models and problems. *Genome* 31:768–777
- Henderson CR (1984) ANOVA, MIVQUE, REML, and ML algorithms for estimation of variances and covariances. In: David HA, David HT (eds) *Statistics: an appraisal*. The Iowa State University Press, Ames, pp 257–280