Interbull Meeting, Ottawa, Canada, August 5 - 6, 1994
Open Session Programme

Estimation of breeding values using testday models

Topic-Session I: The national scene; status reports and current advances in national evaluations

H.H. Swalve
Institut für Tierzucht und Haustiergenetik
Albrecht-Thaer-Weg 3
37075 Göttingen, Germany
e-mail: HSWALVE@GWDG.DE

Introduction

In dairy cattle evaluation procedures models which directly consider records of individual testdays have become of interest. In a broad sense, all models which incorporate records from single testdays can be defined as testday models. This incorporation may consist of a precorrection of individual testday records for fixed effects and then combining these records for the use in evaluation purposes in a second step as is done in Australia (Jones and Goddard, 1990) or may be the direct consideration of testday records in an appropriate one-step evaluation model. Especially interesting and easy-to-use is the model suggested by Ptak and Schaeffer (1993). This model can be viewed as a repeatability animal model in which single testday (TD) records are taken as repeated measurements and factors to model the curve of the lactation are included. The factors used were derived from work by Ali and Schaeffer (1987) who suggested a regression model to describe the curve of the lactation and demonstrated the advantage over other models, including the classical model by Wood (1967).

Ptak and Schaeffer (1993) show that residual variances are smaller when using herd TD as a definition for contemporary groups instead of herd-year-seasons. They emphasize that extension of records could be avoided, cows could be grouped into different contemporary groups within herd according to their stage of lactation like it is actually done on many farms and, furthermore , that the accuracy of evaluations could be increased if heritabilities of testday records were in the range of those for 305-day records.

In a previous report (Swalve, 1993) it was shown that the latter requirement could be met by TD models. Aim of the present study was to compare breeding values estimated under different TD models with and without the inclusion of herd TD as the main fixed effect in the model and with different reductions for the data according to number of testdays considered. Additionally, the relationship between testday yields in testday models and the persistency of lactation production was to be examined.

Material and Methods

Data was supplied by VIT (Agricultural Computing Center, formerly RLN), Verden and covered a coastal region from northern Germany where herdsizes commonly are large by german standards. All calvings were from years 1985 to 1991. Only first lactation records were used. The preparation of the data has been described in Swalve (1993), the data set used here is identical to the set referred to as "Region 2" in the paper cited. The main step of editing consisted of a restriction on size of herds across years, at least 60 lactations were required. This edit was reasoned by the attempt to enable the application of models which consider the herd effect directly instead of grouping of herds in herdclasses as is commonly done in Germany. The structure of the final data set is given in Table 1, the data set is referred to as "full data set". For computational reasons this data set had to be slightly reduced in order to enable the application of a model including herd TD as a definition for contemporary groups. 21 small herds comprising 1226 cows were discarded thus increasing the number of contemporaries per herd TD in the resulting "reduced data set".

Lactation and TD production was expressed in kg per lactation, and per day, respectively. For the analysis of the persistency of lactation production, three different criteria were defined following work by Sölkner and Fuchs (1987): P31 denotes the ratio of milk yield in days 201 - 305 of lactation divided by yield in days 1 - 100; MAME was defined as the maximum production of all TD records per lactation divided by the respective mean, STD simply is the standard deviation of all TD records of a lactation. Under these definitions MAME and STD cover the entire lactation whereas P31 uses production records with standardized lengths.

Variance components were estimated using DFREML 2.1 (Meyer, 1993) applying univariate models. Single trait BLUP evaluations were carried out using the package PEST 3.0 (Groeneveld, 1993). The description of models can be summarized as follows:

Model	Name of model	Short name for EBV[1)]	Trait	Main fixed effects	Covariates[3)]	Random effects
I	TD-HTD[2)]	BV-HTDx	testday yield	herd-testday	age of calving, DIM/c, $(DIM/c)^2$, ln(c/DIM), $(\ln(c/DIM))^2$	animal, permanent environment
II	TD-HYS	BV-TDx, BV-TD8, BV-TD3_7	testday yield	herd-year-season	age of calving, DIM/c, $(DIM/c)^2$, ln(c/DIM), $(\ln(c/DIM))^2$	animal, permanent environment
III	305-HYS	BV-305	305-d yield	herd-year-season	age of calving, DIM1	animal
IV	P-HYS	-	P31, MAME, STD	herd-year-season	age of calving, DIM1	animal

1) Short names for estimated breeding values:
...TDx: use of all available TD records per cow
...TD8: use of only TD records 1 - 8 per cow
...TD3_7: use of only TD records 3 - 7 per cow
2) Applied to reduced data set only
3) DIM = days in milk, DIM1 = interval from calving to first TD

Models I and II are similar to the TD models suggested by Ptak and Schaeffer (1987). For the estimation of breeding values using the full data set under a HYS model (model II), a reduction in the amount of information used per cow was set up so that besides one estimation run using all available information, a run considering only testdays 1 to 8 and a run including only mid-lactation testdays, i.e. TD3 to TD 7, was carried out. This reduction could not be applied under model I (TD-HTD) since the number of contemporaries within a herd TD would have been too small.

Results and Discussion

In Table 2 raw means and standard deviations for lactation and part lactation yields are given along with those for the three criteria of persistency. The averages show typical values found for first lactation yields in northern Germany. Persistency criteria do exhibit values not uncommon for the level of production. However, the variation of the STD criterion appears to be rather high. The distribution of all criteria of persistency was not found to differ significantly from a normal distribution.

Table 3 displays the results from the estimation of variance components. For TD models, only results when using all available testdays for each cow are reported here. Estimates of heritabilities under the TD-HTD model (I) are substantially higher than those found for model II (TD-HYS). This is not only due to a reduction in the residual component but also to an increase in the additive genetic variance. The result for protein yield under model I is somewhat puzzling since a rather high heritability is estimated along with a phenotypic variance lower than could be expected. All heritabilities under the TD-HTD model are in the range of those for model III (305-HYS) thus confirming the preliminary results reported by Swalve (1993). The estimates for the persistency traits in general are in good agreement with the literature (Danell, 1982; Leukkunen, 1985) although somewhat lower than those reported by Sölkner and Fuchs (1987). The STD criterion of persistency appears to exhibit the highest heritability. This is rather surprising - but not unexpected from the literature (e.g. Sölkner and Fuchs, 1987) - since it could be argued that the standard deviation of TD yields would be highly influenced by causes not included in the model and with presumable low heritability of these causes, e.g. sudden changes in diet, diseases, etc.

For the reduced data set, two runs estimating breeding values under models I and III, respectively, were carried out. Correlations between the two sets of estimated breeding values, BV-305 and BV-HTDx, are given in Table 4 along with correlations of these breeding values with phenotypic measurements of persistency for cows. The correlations between breeding values are around .80 for all traits and therefore rather high for breeding values based on first lactations only of cows. These correlations can not be interpreted as estimates of genetic correlations, not even if a correction for the accuracies of the estimated breeding values could be done as has been suggested in studies dealing with different data structure and traits (Calo et al., 1973; Blanchard et al., 1983). Two antagonistic effects hinder an interpretation of these correlations as genetic correlations: Due to low accuracies the values would be corrected upward but due to the fact that both traits were measured on the same animals and therefore should have an environmental covariance, possibly positive, the values would have to be corrected downward.

The correlations between estimated breeding values and persistency given in Table 4 are generally low for both sets of breeding values, favourable for P31 and MAME and unfavourable for STD. A slight tendency for more favourable correlations is observed for HTDx.

Corresponding to Table 4, Table 5 displays shifts in rank of animals sorted by either BV-305 or BV-HTDx. The results are shown for various top lists as an indicator for the situation of a selection of bull dams or sires. As may be expected more drastic changes occur for cows as compared to sires with more than 20 daughters due to lower accuracy of cow evaluation. Largest rank shifts are given based on the ranking for BV-305. For cows substantial shifts in rank occur even when only the top ten cows are considered. The most drastic shift for this list is found for protein yield. Since the rank shifts are given based on BV-305 ranks compared to BV-HTDx ranks, almost all shifts are downward. Only on the sire lists two upward shifts are observed for milk and protein yield for the same sire. In general, rank shifts for sires appear to be rather small.

As explained above, only HYS models could be applied to the full data set. However, for this data set the effect of a reduction in the amount of information used per cow could be analyzed. Table 6 shows the resulting correlations among estimated breeding values along with the correlations of breeding values with the criteria of persistency. All correlations between breeding values are extremely high, in fact close to unity. Compared to the results given in Table 4 this underlines the importance of a presumably positive environmental correlation between the two traits which is not accounted for in the comparison of two univariate estimations of breeding values and here has drastic consequences since an identical definition for the contemporary groups is used. Again, correlations with persistency are generally low and in a favourable direction for P31 and MAME, but in a unfavourable direction for STD. For fat yield, correlations with persistency values seem to be lower than for milk and protein yield.

From the results given in Table 6 no drastic changes in ranking of animals from BV-305 to TD lists could be expected. This is confirmed in Table 7 in which the rank changes are presented analogous to Table 5, however, for milk yield only. Comparing the three TD lists for cows, a slightly lesser extent of rank changes seems to be found for BV-305 vs BV-TDx than for the two TD lists incorporating less information per cow. However, for all maximum rank changes within top list, different animals appear in the three TD lists for cows. For bulls, the changes in rank are negligible.

The shifts in rank from the BV-305 top list to the four different TD top lists including the analysis from the reduced data set were used as variables for further analysis. A shift variable was defined as the deviation of rank (BV-305 top list) - rank (TD top list). The results for correlations of these variables with the criteria of persistency is given in Table 8. From the analysis of the full data set it appears that the correlations for the rank deviation BV-305 - BV-TD8 are different from the two other sets of correlations for BV-305 - BV-TDx and BV-305 - BV-TD3_7. For rank shifts from the 305 list to the TD8 list a correlation of -.33 is exibited with P31. This essentially means that the more negative the rank shift deviation is, the more persistent is the lactation production over testdays. A negative rank shift deviation can only be found if the rank on the TD list is lower (higher numerical value) than the one on the 305 list, i.e. there appears to be a tendency for more persistent lactation production among cows that rank high on the BV-305 list as compared to the BV-TD8 list. However, the correlation for MAME is close to zero and the correlation for STD is .12, i.e. a slight tendency for increased standard deviations of TD production can be observed for cows that rank lower on the BV-305 list than on the BV-TD8 list.

For the deviation of rank(BV-305) - rank(BV-TDx) the picture looks different. The correlation with P31 is now reduced to -.17, i.e. it shifted in a direction more favourable for the TD model and the correlations for MAME and STD are both negative, i.e. favourable for the TD model since a tendency for a reduction in the MAME and STD values is observed for cows ranking higher on the TD list (positive rank deviation). An explanation for this appears to be that for TDx all available information per cow is used thus leading to a more favourable relationship with persistency. A graphical representation of the two correlations with P31 for shifts BV-305 to BV-TDx and BV-305 to BV-TD8 showing all values of the 15,756 observations is given in Figures 1 and 2.

Unexpectedly, the correlations for BV-TD3_7, with the exception for STD, are very similar to those for BV-TDx. With respect to the persistency of lactation production a focus on the midst of lactation seems to be as advantageous as considering the entire lactation whereas a cut-off at the eighth TD seems to be more detrimental for the persistency if TD models are used for evaluation and selection purposes. In the first row of Table 8, the correlations for rank shifts from BV-305 to BV-HTDx are given. These are very low and favourable for MAME and STD with respect to the application of the HTDx model.

Conclusions

For an analysis of the potential changes in ranking of animals when switching from traditional 305-day models to testday models models similar to those proposed by Ptak and Schaeffer (1993) were used. Models considering herd-testday (HTD) as a definition for contemporary groups seem to be advantageous with respect to the variance components estimated. Under HTD models substantially higher heritabilities can be found for testday records as compared to models that use a more traditional herd-year-season classification. However, such models require relatively high sizes of herds and are computationally demanding. Shifts in rank of animals sorted by estimated breeding values for HTD and 305-day models indicate only minor changes for sires but rather drastic ones for cows. In general, testday models do not seem to have a negative effect on the relationship between production yields and the persistency of the lactation production as long as either the entire lactation or only the midst of lactation is considered.

References

Ali, T.E., and L.R. Schaeffer (1987): Can. J. Anim. Sci. 67: 637
Blanchard,P.J., R.W. Everett, and S.R. Searle (1983): J. Dairy Sci. 66: 1947
Calo, L.L., R.E. McDowell, L.D. Van Vleck, and P.D. Miller (1973): J. Dairy Sci. 37: 676
Danell, B. (1982): Acta Agric. Scand. 32 : 93
Groeneveld, E. (1993): PEST user's manual for version 3.0. Mimeo.
Jones, L.P., and M.E. Goddard (1990): Proc. 4th World Congr. Genet. Appl. Livest. Prod., Edinburgh, UK, XIII:382.
Ptak, E., and L.R. Schaeffer (1993): Livest. Prod. Sci. 34: 23
Leukkunen, A. (1985): Z. Tierzüchtg. Züchtgsbiol. 102 : 117
Meyer, K. (1993): DFREML 2.1 Program Package and User Notes, AGBU, Armidale, Australia
Sölkner, J., and Fuchs, W. (1987): Livest. Prod. Sci. 16 : 305
Swalve, H.H. (1993): Proc. Interbull Meeting, Aarhus, Denmark
Wood, P.D.P. (1967): Nature 216: 164

Table 1: Structure of the data

	Full data set	Reduced data set
No. of 1. lactations (with ≥ 8 testday records)	15,756	14,530
No. of testday records	155,494	143,532
No. of herds	215	194
No. of sires	783	763
No. of herd-year-seasons	2660	2400
No. of herd-testdays	17,568	14,984
Average no. of testday records per cow	9.86	9.88
Average no. of testday records per herd[1]	723.2	739.9
Minimum no. of testday records per herd[1]	524	588
Maximum no. of testday records per herd[1]	1293	1293

[1] across years 1985 to 1991

Table 4: Correlations between estimated breeding values from TD and 305-d models (BV-HTDx, BV-305) and phenotypic measurements of persistency for cows[1)]

	Estimated breeding value	Correlation between EBVs	P31	MAME	STD
Milk yield	BV-HTDx	.82	.19	-.16	.17
	BV-305		.23	-.15	.22
Fat yield	BV-HTDx	.82	.11	-.07	.20
	BV-305		.14	-.03	.23
Protein yield	BV-HTDx	.79	.21	-.18	.10
	BV-305		.22	-.12	.15

[1)] Correlations between EBVs for bulls were: (all bulls) .82, .83, .80, and (only bulls ≥ 20 daughters) .85, .85,.84, for milk, fat, and protein yield, respectively.

Table 2: Raw means and standard deviations of first lactation yield and persistency traits (full data set)

Trait	Definition of trait	Milk yield		Fat yield		Protein yield	
		mean	std.dev.	mean	std.dev.	mean	std.dev.
L305	305-day lactation yield	6007	1007	251.2	44.4	198.7	32.6
P100	Yield day 1 - 100	2316	363	95.0	17.3	73.2	11.2
P200	Yield day 101 - 200	1990	365	82.8	15.5	66.4	12.0
P305	Yield day 201 - 305	1700	393	73.4	17.2	59.2	13.6
P31	Persistency: (yield day 201 -305) / (yield day 1 - 100)	.74	.14	.78	.16	.81	.16
MAME	Persistency: (Maximum of TD yields) / (Mean of TD yields)	1.29	.12	1.31	.16	1.25	.12
STD	Persistency: St. dev. of TD yields	3.81	1.33	.15	.06	.11	.04

Table 3: Estimates of variance components, heritabilities, and effect of permanent environment (p.e., for TD-models only) for testday and 305-day milk yields and persistency of lactation.

	σ^2_A	σ^2_E	σ^2_{PE}	h^2	p.e.
Analysis of testday records applying testday models I and II					
Model I, TD-HTD					
Milk yield	9.00	4.11	10.92	.37	.45
Fat yield	.0145	.0112	.0219	.30	.46
Protein yield	.0080	.0051	.0055	.43	.30
Model II, TD-HYS					
Milk yield	7.47	5.69	9.59	.33	.42
Fat yield	.0122	.0143	.0213	.26	.46
Protein yield	.0065	.0073	.0115	.26	.45
Analysis of 305-day records applying model III					
Milk yield	254,004	397,920	-	.39	-
Fat yield	350.8	745.0	-	.32	-
Protein yield	175.4	407.4	-	.30	-
Analysis of persistency applying model IV					
Milk yield: P31	.0015	.0130	-	.11	-
MAME	.0016	.0095	-	.14	-
STD	.1906	1.0862	-	.15	-
Fat yield: P31	.0017	.0160	-	.10	-
MAME	.0014	.0186	-	.07	-
STD	.0003	.0024	-	.11	-
Protein yield: P31	.0014	.0167	-	.08	-
MAME	.0007	.0095	-	.07	-
STD	.0001	.0010	-	.10	-

Table 5: Demonstration of rank shifts based on animals ranked by BV-305 (305-list) compared to ranking by BV-HTDx (TD-list) in various top lists

List considers … animals	Milk yield			Fat yield			Protein yield		
	No. of animals on both lists	Largest rank shift		No. of animals on both lists	Largest rank shift		No. of animals on both lists	Largest rank shift	
		From 305-list #	To TD-list #		From 305-list #	To TD-list #		From 305-list #	To TD-list #
Cows with records (n=14,530)									
first 10	3	5	145	5	3	100	4	9	501
first 20	8	17	595	10	20	165	6	16	777
first 50	22	22	710	20	33	725	24	43	1143
first 100	44	96	1711	50	59	1070	44	90	1213
Bulls (≥ 20 daughters, n=107)									
first 5	2	4	11	4	5	15	3	2	7
first 10	7	8	15	8	6	16	7	8	21
first 20	16	19	35	14	14	28	15	12	35
first 50	46	45	2	43	48	80	41	50	8

Table 6: Correlations between estimated breeding values from TD and 305-d models (BV-TDx, BV-TD8, BV-TD3_7, BV-305) and phenotypic measurements of persistency for cows

	Estimated breeding values				Persistency		
	BV-TDx	BV-TD8	BV-TD3_7	BV-305	P31	MAME	STD
BV-TDx							
Milk yield	*	.99	.97	.99	.21	-.17	.19
Fat yield	*	.98	.95	.98	.11	-.06	.20
Protein yield	*	.98	.95	.98	.19	-.15	.12
BV-TD8							
Milk yield		*	.98	.98	.18	-.14	.24
Fat yield		*	.95	.98	.05	-.02	.27
Protein yield		*	.97	.94	.13	-.11	.19
BV-TD3_7							
Milk yield			*	.97	.19	-.17	.21
Fat yield			*	.94	.13	-.13	.19
Protein yield			*	.94	.17	-.18	.15
BV-305							
Milk yield				*	.23	-.14	.22
Fat yield				*	.14	-.03	.23
Protein yield				*	.22	-.12	.16

Table 8: Correlations between phenotypic measurements of persistency and rank shifts from 305-day top list (BV-305) to four different testday top lists (BV-HTDx, BV-TDx, BV-TD8, BV-TD3_7) for cows (milk yield)

Rank (R) deviation 305-list - TD-list	P31	MAME	STD
R(BV-305) - R(BV-HTDx)	-.04	-.07	-.07
R(BV-305) - R(BV-TDx)	-.17	-.21	-.12
R(BV-305) - R(BV-TD8)	-.33	-.06	.12
R(BV-305) - R(BV-TD3_7)	-.20	-.21	-.02

Table 7: Demonstration of rank shifts based on animals ranked by BV-305 (305-list) compared to ranking by testday model breeding values BV-TDx, BV-TD8, and BV-TD3_7 (TD-list) in various top lists for milk yield

List considers … animals	BV-TDx			BV-TD8			BV-TD3_7		
	No. of animals on both lists	Largest rank shift		No. of animals on both lists	Largest rank shift		No. of animals on both lists	Largest rank shift	
		From 305-list #	To TD-list #		From 305-list #	To TD-list #		From 305-list #	To TD-list #
Cows with records (n=15,756)									
first 10	7	6	13	6	10	37	5	7	35
first 20	15	18	43	17	10	37	14	16	45
first 50	40	38	144	41	39	116	37	39	96
first 100	79	62	182	84	88	370	77	88	317
Bulls (≥ 20 daughters, n=113)									
first 5	4	5	7	4	5	8	5	-	-
first 10	10	8	5	9	8	3	10	7	10
first 20	18	17	23	18	20	27	18	20	28
first 50	49	47	37	49	47	36	48	48	64

Figure 1: Relationship between persistency P31 and rank shift deviation R(BV-305) - R(BV-TDx), r = -.17

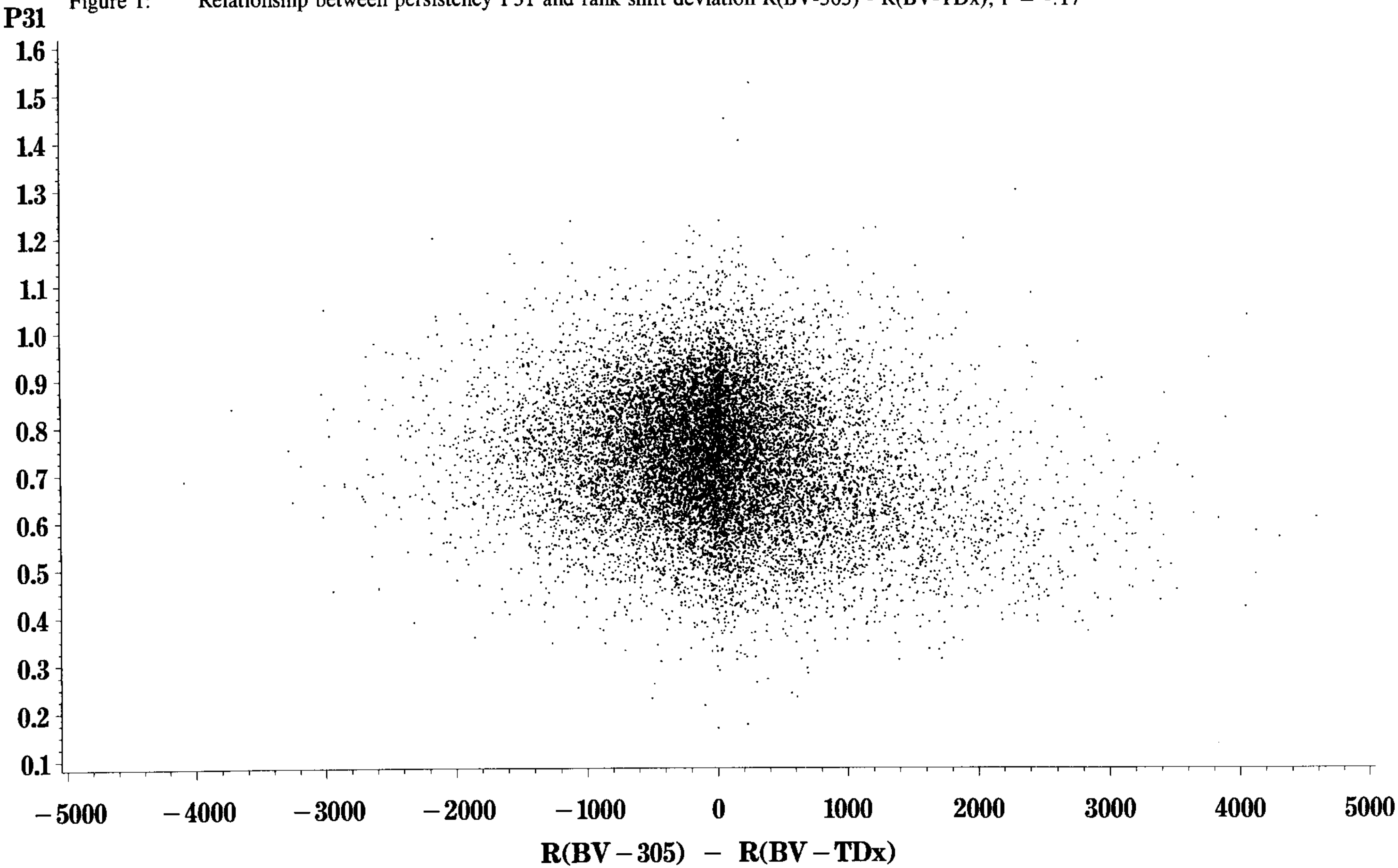

Figure 2: Relationship between persistency P31 and rank shift deviation R(BV-305) - R(BV-TD8), r = -.33