UDK 636.082.2

Izlaganje sa znanstvenog skupa
Conference paper

USE OF TEST DAY RECORDS FOR GENETIC EVALUATION

H. H. Swalve

Summary

Test day records conventionally have been used in aggregated forms as lactation records. Test day models are more precise since compared to traditional models changes of the environment during lactation can be included. In general, three approaches exist: Correction of environmental effects on the test day level and later processing of lactation records in a conventional way, test day models that directly consider records within a lactation as repeated traits and account for the curvilinear pattern of the lactation curve by a suitable sub-model, and random regression models that include the latter features but additionally decompose the animal's effect into random regression coefficients. Random regression models are related to the general approach of covariance functions proposed for longitudinal data. Due to their flexibility, their precision and their further potential to also provide management aids test day models will be the method of choice in future evaluation procedures for dairy animals if not already implemented.

Keywords: test day records, genetic evaluation, lactation curve

Introduction

Records gathered at so-called test days are the basis of dairy animal (cattle, sheep, goat) recording systems. Recently, attention has been drawn to the use of test day records as they are recorded instead of the use of lactation records that are calculated from them. Reasons for this are: 1. Conventional recording systems are costly, thus calling for an extensification of data collection

Rad je priopćen na "6th World Congress on Genetics Applied to Livestock Production", Armidale, 1998.

H. H. Swalve, Institute of Animal Breeding and Genetics, University of Göttingen, Albrecht-Thaer-Weg 3, 37075 Göttingen, Germany

schemes. Such extensifications could be prolonged intervals than those used today, alternate recording (a.m./p.m. schemes), or mixtures between owner-sampler schemes and official protocols. Such measures would it make more difficult to calculate lactational or yearly records thus raising the question of the value of single test day records. 2. In order to reduce generation intervals, evaluations are sought that can rapidly make use of every single test day record that is available at a given time. 3. The traditional approach of using lactation records as been criticized as inconsistent since records taken at defined locations and time are aggregated in a rather trivial way and are subsequently subject to quite sophisticated statistical analyses targeting for an optimum differentiation of all genetic and environmental effects.

With test day models it is attempted to account for systematic, environmental and genetic effects directly there, where they are expressed, at the day of recording. The aim is to give alternative formulations of selection criteria. Under test day models, no assumption about or agreement on the "normal" length of a lactation has to be made. Test day models are more flexible. Records do not have to be extended using some factors and regulations, instead, every piece of information can be used.

Test day models have to be differentiated from models describing the typical curves of dairy production traits without an initial intention to separate genetic and environmental effects. However, many test day models quite naturally make use of these functions and thus lactation curve functions could be called sub-models from the test day model point of view. A classical example of work on lactation curve functions is the study by Wood (1967), an excellent review can be found in Masselin et al., (1987), later work also exists (e.g. Guo and Swalve, 1995).

Genetic parameters for test day records

A test day model (TDM) denotes a model considering several test days per individual lactation. According to this definition, a model to analyze single test day records separately is not a test day model. However, some results from genetic analyses of individual test day records shall be reported here to give some insight in what possibly can be expected from test day models. A more detailed compilation, also including the modeling of environmental effects can be found in Swalve (1995a). Table 1 displays literature results from genetic analyses of the past 10 years considering test day records individually. The general pattern of the heritability estimates shows an increase of the magnitude towards mid-lactation and again a slight decrease at the end of lactation. This pattern is also visible in an even more pronounced form for the genetic correlation between individual test days and 305-day lactation records with the results from Rekaya et al., (1995) as an exception. The magnitude of heritability estimates of individual test days never reaches those for the corresponding ones for 305-day records but gets close to them in mid-lactation. From this, a use of evaluating individual test day records jointly can already be derived since more information modeled more precisely can be combined. The question to raise here is whether all this information should be and how it should be combined. In studies that estimated the genetic and phenotypic correlations among test day records applying multiple trait models (e.g. Meyer et al., 1989; Pander et al., 1992; Rekaya et al., 1995) it has been shown that fairly close genetic relationships exist between adjacent test days. Correlations drop for test days that are far apart. Relationships among test days in mid-lactation usually are highest and close to unity. These results thus suggest that it may not be very useful to consider all individual test days jointly in a multiple trait evaluation with it's obvious advantages but also inherent problems of large (co)variance matrices.

Strategies for test day models

In recent years, different strategies have been proposed for the use of test day records in genetic evaluations. In general, two approaches may be differentiated: The use of test day records in a one-step method directly producing breeding values for dairy production, and two-step methods that apply some correction for environmental effects on the test day level but run evaluations on records or residuals combined after this first step. A more detailed differentiation of methods may be the following:

2-step test day model (2STDM). The general form of this approach has been suggested and implemented in Australia (Jones and Goddard, 1990), New Zealand (Johnson, 1996) and the North-Eastern part of the USA (Everett et al., 1994; ABC, 1997).

Table 1. - ESTIMATES OF HERITABILITIES (X 100)AND GENETIC CORRELATIONS (X 100) OF TEST DAY RECORDS WITH 305-DAY FIRST

Author	Method[A]	Definition of records[B]	Parameter	Records of individual test days										305-day record
				1	2	3	4	5	6	7	8	9	10	
Wilmink,	REML	CSI	h^2	16	25	29	30	30	29	29	25	26	-	31
(1987)	SM		$-r_g$	61	88	94	95	99	99	95	99	80	-	-
Meyer et al.	REML	TDI	h^2	20	22	25	27	24	25	24	24	21	17	37
(1989)[C]	SM													
Pander et al.	REML	TD	h^2	22	33	34	36	35	38	39	43	-	-	49
(1992)[D]	SM		r_g	87	89	97	98	99	97	98	97			
Reents et al.	REML	TD	h^2	10	14	21	30	32	37	35	31	30		
(1994)	AM													
Swalve	REML	TD	h^2	18	24	28	33	33	36	31	26	-	-	39
(1995b)	AM	ASI	h^2	18	24	29	34	34	37	35	33	30		
Rekaya et al.	REML	TD	h^2	18	19	22	21	23	25	25	27	28	27	31
(1995)	AM		r_g	89	76	72	74	78	74	70	72	68	72	

[A] REML = Restricted Maximum likelihood; SM = sire model; AM = animal model

[B] TD = test day record as recorded with given days in milk; TDI = test day records within equidistant intervals; CSI = cumulated yield within a 30-day interval; ASI = average yield within a 30- day interval

[C] Results pooled from estimates of three data sets as given in the study; complete lactation here is 300 days, heritability given for the lactation record is the average from three data sets from bivariate analyses

[D] Results given for data set 1 which was used for simultaneous analysis of TDI1 to TD8

Under the Australian system that was implemented in 1984 (Jones and Goddard, 1990), test day records have been adjusted before combining them into lactation records. Johnson (1996) reports that New Zealand had used a test day approach for the evaluation of cows. Under the newly developed system estimates for the correction of HTD effects are calculated applying an extended version of Wood's model (Wood, 1967). Test day records, or rather their residuals after correction are then combined into lactation records weighting the individual test day record according to the correlations among them. The standardized lactation records are then subject to a further analysis using an animal model. The Cornell group (ABC, 1997) adjusts test day records for age, DIM, month of calving and progeny effects within herd before combining them into 320-day lactation records.

Ordinary test day models (OTDM). Based on the work of Beard (1983), Meyer et al., (1987) applied a model fitting herd-test-day, sire of cow and cow along with covariables accounting for age at calving and DIM to obtain REML estimates of variance components. Test day records were taken as repeated measurements. Ptak and Schaeffer (1993) proposed a rather similar approach for an animal model. Under their model the effect of DIM is accounted for by a set of regression coefficients corresponding to the lactation curve function suggested by Ali and Schaeffer (1987). Ptak and Schaeffer (1993) emphasize, that the heterogeneity of the residual variance of test day records in the course of the lactation should be considered in such a model. Furthermore, the regressions should be nested within age-season and possibly other effects to allow for different curves for distinct groups of cows. Modeling the herd effect as HTD instead of herd-year-season of calving lead to reduced residual variances. This result has subsequently been confirmed by several studies applying the Ptak and Schaeffer (1993) model (Swalve, 1995b; Pösö et al., 1996; Strabel and Szwaczkowski, 1997). Reents et al., (1995ab) extended this approach to a multiple trait model that considers the first three lactations as different traits, within a lactation records are taken as repeated measurements. Originally, their approach was developed for the evaluation of somatic cell score (SCS). Meanwhile, this model has been used for the national evaluations for SCS in Canada and Germany (Reents et al., 1995c; Reents, 1996).

Covariance Functions and Random Regression test day models (RRTDM). Covariance functions are the equivalent of covariance matrices for traits with many, potentially infinitely many, records in which the covariances are defined as a function of age or time (Meyer and Hill, 1997). Starting from a multiple trait approach, covariance functions offer the opportunity to reduce the rank of the covariance matrix among (highly) correlated traits. Orthogonal polynomials are a suitable family of functions to describe covariance functions. The estimation of the matrix of the coefficients of the covariance function has been demonstrated Kirkpatrick et al., (1990, 1994). This approach is related to earlier work that dealt with finding the number of independent combinations of correlated traits as combinations of the largest eigenvalues of covariance matrices. Wiggans and Goddard (1997) report that instead of using the method of finding the largest eigenvalues a definition of the linear combination of traits (test day records) according to functions derived from the selection objective would preferable. Starting from a model with 60 traits (milk, fat, protein, two parities, 10 test days per lactation) they were able to reduce the rank of the covariance matrix to an order of six. For the six 'new' traits a canonical decomposition is feasible thus reducing the computing demand. The method of Wiggans and Goddard can also be viewed as a 2STDM since some effects are corrected within herd and the resulting records are subject to an across-herd analysis.

According to Meyer and Hill (1997) it can be shown that the covariance function model is equivalent to the random regression model (RRM). In the RRM, regression coefficients are taken as random thus allowing for a covariance structure among them (Henderson Jr., 1982). The use of RRM for test day records was suggested by Schaeffer and Dekkers (1994) and subsequently has been successfully applied to estimate genetic parameters and for the genetic evaluation (Jamrozik et al., 1996; Jamrozik and Schaeffer, 1997; Jamrozik et al., 1997ab). Jambrozik and Schaeffer (1997) define a model based on the OTDM by Ptak and Schaeffer (1993). The sub-model to account for the curve of the lactation is used in a fixed form (nested in province-age-season) and a random form thus replacing the animal's genetic effect by five regression coefficients (intercept plus four parameters). They show that by this formulation the genetic variance can be calculated for every test day or any combination of test days during the lactation.

Genetic parameters under test day models

Heritabilities estimated under the three strategies as summarized above are given in Table 2. Under the 2-step approach, the magnitude of heritabilities can be compared directly to the corresponding values from 305-day records. All three studies applying a 2STDM found higher estimates for the test day approach. This clearly indicates an advantage of TDM with respect to accuracy of genetic evaluations. For 1-step approaches, such an advantage is not obvious when comparing both sets of estimates unless the estimates for the

TDM are of equal or greater magnitude compared to the traditional ones. This quite drastically is the case for the study of Strabel and Swaczkowski (1997), the other authors found smaller heritabilities for TDM. For those studies the accuracy of an evaluation of individual animals would have to be calculated according to both approaches in order to consider the heritability jointly with the number of records that go into the evaluation. A rough comparison of OTDM and RRTDM estimates may be made using the results from Reents et al., (1995a) and Jamrozik and Schaeffer, (1997) since both use Canadian data. This indicates slightly higher heritabilities under the RRTDM. For both studies the officially used Canadian heritabilities of .33 for milk, fat, and protein yield (Jambrozik and Schaeffer, 1997) provide a comparison with the traditional approach thus indicating that accuracy of evaluations should increase when using a TDM.

Topics of ongoing and future research

Choice of a sub-model. In general, two different approaches exist. The first method is to use variance functions as suggested by Meyer and Hill (1997). This could be called a method in which 'the data decides on the function to be used'. A different approach would be to decide on a function to describe the lactation curve a-priori as is done in OTDM. Guo and Swalve (1995) evaluated eight functions with a different number of parameters (from three to six). Goodness of fit was analyzed using data from daily milk recording as is done with modern milking parlor equipment. In general, a better fit was obtained by functions with more parameters. Within each group of functions differences were small.

The question arises, which criteria should be used for comparisons. As Ali and Schaeffer (1987) point out, some functions may do well in predicting 305-day yields while others are better in predicting daily yields. Jamrozik et al., (1997b) compared two different functions to be used in a RRTDM. They suggest that for the fixed set of regressions functions different from those used in the random part could be used. However, they found a slight advantage with respect to prediction error variance for their model used previousy (Jamrozik and Schaeffer, 1997) that had five parameters for the fixed and the random part. Jamrozik et al., (1997b) point out that a compromise might have to be found in order to reduce the number of parameters and thus reducing the computational demand which is substantial under a RRTDM.

Table 2. - ESTIMATES OF HERITABILITIES FOR MILK YIELD FROM TEST DAY APPROACHES (2STDM, OTDM, RRTDM)

Author	Method	Test day approach			Corresponding 305-d estimates		
		Milk	Fat	Protein	Milk	Fat	Protein
Test day combined into lactation records after precorrection (2STDM)							
Meyer et al., (1989)[A]	Multiple trait	38	-	-	36		
	REML, SM	42	-	-	42		
		33	-	-	33		
Van Tassell et al., (1992)[B]	Multiple trait REML	+12%	+11%	+17%			
Visscher and Goddard (1995)[C]	Multiple trait	33	29	29	24	20	21
	REML, SM	39	37	35	28	26	26
Ordinary test day repeatability models (OTDM)							
Meyers et al., (1992)[D]	REML, SM	17	14	13	17	13	
Swalve (1995a)[E]	REML, AM	24	16	16	39	32	30
		28	18	19			
Reents et al., (1995)[F]	Multiple traits	30	25	25			
	Bayes (Gibbs), AM	24	20	23			
		23	18	20			
Pösö et al., (1996)[G]	REML, SM	23	-	-	32		
Strabel and Swaczkowski (1997)[H]	REML, AM	27	22	25	16	10	12
		29	20	20	15	10	11
Test day repeatability models with random regression (RRTDM)							
Jambrozik and Shaeffer (1997)	Bayes (Gibbs), AM	32	28	28			

[A] Results from three data sets, 9 test days per lactation considered
[B] Estimates are not given, rather the per cent increase compared to the 305-d approach
[C] First row Holstein Friesian, second row Jersey
[D] Corresponding estimates taken from Mayer (1985) for data from NSW that was also used in this study
[E] First row herd-year-season model, second row herd-test-day model
[F] Estimates for first, second, and third lactation
[G] Model with herd-test-day effect
[H] First row raw records, second row transformed (Box-Cox) records

Heterogeneity of variance. Ptak and Schaeffer (1993) and Jamrozik and Schaeffer (1997) emphasize that the heterogeneity of the residual variance in the course of the lactation should be accounted for. In their application of a RRTDM (Jamrozik et al., 1997a) they allow the residual variance to differ from day to day. The problem of heterogeneous variances is a general one, genetic and environmental variances differ among subgroups of data. Methods to account for this phenomenon (e.g. Wiggans and Van Raden, 1991) need to be refined for their use in test day models.

Multiple trait models. One argument in favor of multiple trait test day models comprising milk, fat and protein is that also data from test days may be used where only yield of milk is recorded (Wiggans and Goddard, 1997; Jamrozik et al., 1997). Other multiple trait approaches consider performance in different lactations as different traits. Other traits could be included, notably somatic cell score or even more traits that are linked to test days. Computationally, this will be difficult to achieve on a national basis given today's tools for computing. However, advances in hardware and software are rapid and thus even a RRTDM comprising several traits lactations may not be far from a large-scale national implementation.

Connclusion

Test day models give a more precise fit for the dairy animal recording situation than other models. Contemporary groups could even be defined within-herd according to management groups, frequency of milking, treatment with performance-promoting substances, etc. Additionally, accuracy of evaluations for individual animals is expected to increase since more, data can be used. Test day models are more flexible than traditional 305-day approaches. This is of special importance for mixed and potentially more extensive recording schemes (longer intervals, missing tests for fat and protein, etc.), every piece of information can be used. Allowing for individual variation of the lactation curve as done in random regression methods is a step towards a more biological approach of dairy performance. Test day models naturally would call for frequent evaluation runs (Jamrozik et al., 1997a), as new information comes into the data storage system it could be used immediately. This not only could have an impact on reducing the generation interval but side-products of evaluations also could be used as a management aid to dairy farmers (Pösö et al., 1996). Ultimately, the dairy farmers themselves have to pay for evaluations, so it should be obvious that their investment is used for genetic and environmental improvement of their dairy business.

REFERENCES

1. ABC (1997): Animal Breeding Cornell http://www.abc.cornell.edu/abc/siretabs/about.html.
2. Ali T. E., L. R. Schaeffer (1987): Can. J. Anim. Sci. 67: 637-644.
3. Beard, K. (1983): M. sc. thesis, University of Melbourne.
4. Everett, R. W., F. Schmitz, L. H. Wadell (1994): J. Dairy Sci. 77 (Suppl. I): 267
5. Guo, Z., H. H. Swalve (1995): Interbull Mtg. Prague, Uppsala, Interbull Bull. No.11.
6. Henderson, C. R. Jr. (1982): Biometrics 38: 623-640.
7. Jamrozik, J., L. R. Schaeffer, J. C. M. Dekkers (1996): Interbull Mtg. Veldhoven, Uppsala, Interbull Bull. No.14:124-134.
8. Jamrozik, J., L. R. Schaeffer (1997): J. Dairy Sci. 80: 762-770.
9. Jamrozik, J., L. R. Schaeffer, J. C. M. Dekkers (1997a): J. Dairy Sci. 80 (accepted).
10. Jamrozik, J., G. J. Kistemaker, J. C. M. Dekkers, L. R. Schaeffer (1997b): J. Dairy Sci. 80 (accepted).
11. Johnson, D. L. (1996): Proc. N. Z. Soc. Anim. Prod 56:16-18.
12. Jones, L. P., M. E. Goddard (1990): Proc. 4. WCGALP, Edinburgh, Vol 13: 382-385.
13. Kirkpatrick, M., D. Lofsvold, M. Bulmer (1990): Genetics 124: 979-993.
14. Kirkpatrick, M., W. G. Hill, R. Thompson (1994): Genet. Res. 64: 57-69.
15. Masselin, S., D. Sauvant, P. Chapoutot, D. Milan (1987): Ann. Zootech. 36:171-206.
16. Meyer, K., W. G. Hill (1997): Livest. Prod. Sci. 47:185-200.
17. Meyer, K., H. U. Graser, S. P. Smith (1987): Report, AGBU, UNE, Armidale.
18. Meyer, K., H. U. Graser, K. Hammond (1989): Livest. Prod Sci. 21:177-199.
19. Pander, B.L., W. G. Hill, R. Thompson (1992): Anim. Prod. 55:11-21.
20. Pösö, J., E. A. Mäntysaari, A. Kettunen (1996): Interbull Mtg. Veldhoven, Uppsala, Interbull Bull. No.14: 124-134.
21. Ptak, E., L. R. Schaeffer (1993): Livest. Prod Sci. 34: 23-34.
22. Reents, R. (1996): Interbull Mtg. Veldhoven, Uppsala, Interbull Bull. No.12: 234-242.
23. Reents, R., J. C. M. Dekkers, L. R. Schaeffer (1994): Proc. 5. WCGALP, Guelph, Vol.17: 120-122.
24. Reents, R., J. Jamrozik, L. R.. Schaeffer, J. C. M. Dekkers (1995a): J. Dairy Sci. 78: 2847- 2857.
25. Reents, R., J. C. M. Dekkers, L. R. Schaeffer (1995b): J. Dairy Sci. 78: 2858-2870.
26. Reents, R., J. Jamrozik, L. R. Schaeffer, J. C. M. Dekkers (1995c): Interbull Mtg. Prague, Uppsala, Interbull Bull. No.11.
27. Rekaya, R., F. Bejar, M. J. Carabano, R. Alenda (1995): Interbull Mtg. Prague, Uppsala, Interbull Bull. No.11.
28. Shaeffer, L. R., J. C. M. Dekkers (1994): Proc. 5. WCGALP, Guelph, Vol.18: 443-446.
29. Strabel, T., T. Swaczkowski (1997): Livest. Prod. Sci. 48: 91- 98.
30. Swalve, H. H. (1995a): Arch. Tierz. 38: 591-612.

31. Swalve, H. H. (1995b): J. Dairy Sci. 78: 929-938.
32. Visscher, P. M., M. E. Goddard (1995): J. Dairy Sci. 78: 205-220.
33. Wilmink, J. B. M. (1987): Doctor's thesis, Wageningen.
34. Wiggans, G. R., P. M. VanRaden (1991): J. Dairy Sci. 74: 4350-4357.
35. Wiggans, G. R., M. E. Goddard (1997): J. Dairy Sci. 80 (accepted).
36. Wood, P. D. P. (1967): Nature 216:164-165.

UPOTREBA ZAPISA O DANU TESTA ZA GENETSKU PROCJENU

Sažetak

Zapisi o danu testa obično se upotrebljavaju kao dio zapisa o laktacijama. Modeli dana testa su točniji jer se u usporedbi s tradicionalnim modelima mogu uključiti promjene okoline za vrijeme laktacije. Općenito postoje tri pristupa: korekcija djelovanja okoline na razinu dana testa i kasnije obrađivanje zapisa o laktaciji na uobičajeni način, modeli dana testa koji izravno uzimaju u obzir zapise u laktaciji kao ponavljanje osobine te daju krivuljast uzorak laktacijske krivulje odgovarajućim sub-modelom, te slučajni modeli regresije što uključuje ove značajke ali dodatno raščlanjuju djelovanje životinje u slučaju koeficijentne regresije. Modeli slučajne regresije u vezi su s općim pristupom funkcija kovarijance predloženih za longitudinalne podatke. Zbog njihovih fleksibilnosti, njihove preciznosti te njihovog potencijala pružanja pomoći upravljanju modelima dana testa bit će metoda izbora u budućim postupcima ocjenjivanja mliječnih životinja, ako već nisu.

Primljeno: 20. 11. 1998.