CHAPTER 1

LITERATURE STUDY

1.1 INTRODUCTION

Achieving genetic improvement involves the identification of those animals with the best breeding values and then ensuring that the selected individuals become parents of the next generation. Artificial insemination (AI) can allow the rapid dissemination of this improvement throughout a population or breed (Bichard, 2002). The importance of accurately estimating breeding values can therefore never be overemphasized.

Dairy breeding has expanded immensely since the mid 1950's, mainly due to the freezing of bull semen without a significant decline in fertility and the consequent development of the AI industry, as well as the development of high-speed computers that enabled genetic evaluation of dairy cattle. Improvement in genetic potential as well as feeding and management practices, resulted in phenomenal increases in milk production of cows. Today breeding of dairy cattle is a highly specialized science, which can be advantageously used by breeders to increase profit within the herds.

Genetic evaluation of dairy sires in South Africa was initially based on progeny groups, using data recorded by the National Livestock Improvement Scheme. This led to the utilization of contemporary comparison methods to estimate the breeding values of sires. Dairy animals in South Africa received breeding values from BLUP (Best Linear Unbiased Prediction) methodology for the first time in 1987, when a Sire Model was implemented for the estimation of breeding values. In 1992 the Animal Model was fitted to dairy records to estimate breeding values in single trait analyses. Since 1999 multitrait analyses were developed for the South African dairy breeds, where completed 305-day milk, butterfat and protein first lactation yields were evaluated together in order to utilize the genetic correlations between the traits for more accurate estimation of these breeding values (Loubser *et al*., 2001). South Africa entered the global arena when genetic groups were incorporated into the pedigrees in order to qualify in 2000 for participation in INTERBULL (International Bull Evaluation Service) runs for the estimation of MACE (Multiple Across Country Evaluation) breeding values. The introduction of the dairy management system IRIS in 2002/2003 in South Africa enabled the extension of incomplete lactation records so that records in progress could be included for the first time in August 2003 in South Africa's genetic evaluations.

Four dairy breeds in South Africa are subjected to genetic evaluation done by the Quantitative Genetics Division of the ARC's Animal Improvement Institute, namely the Ayrshire, Guernsey, Holstein and Jersey breeds. The number of dairy cows participating in milk recording in South Africa, however, does not compare favourably with countries in the rest of the world, as a mere 20% of female animals are officially recorded (Van der Westhuizen, 2002).

In Table 1.1 the number of cows of the different breeds and 240-300 day phenotypic averages for all lactations in the 2004 test year, are indicated (Moore, 2004). The Holstein breed is by far the largest breed participating in Milk Recording in South Africa, accounting to 57% of all dairy cows, followed by the Jersey breed (38%), Ayrshire breed (4%) and the Guernsey breed (1%). It is also interesting to note that almost as many commercial as registered Holsteins and more commercial than registered Guernsey cows participated in Milk Recording during the 2004 test year. As mentioned earlier, only first lactation records were used for genetic evaluation of dairy cows in South Africa, thereby excluding a vast amount of information.

Table 1.1 Number of cows and 240-300 day phenotypic averages (kg/lactation) for breeds participating in performance testing in South Africa during the 2004 test year.

Breed	Number		Milk		BF		Prot		BF %		Prot %	
	R	C	R	C	R	C	R	C	R	C	R	C
Ayrshire	4861	884	6320	5403	255	225	209	179	4.03	4.17	3.30	3.31
Guernsey	588	732	6054	5187	265	230	210	181	4.37	4.44	3.47	3.49
Holstein	39093	33824	8676	6861	329	264	277	220	3.79	3.85	3.19	3.20
Jersey	35701	13063	5455	4809	258	226	202	178	4.72	4.69	3.71	3.70

R=Registered C=Commercial BF=Butterfat Prot=Protein

In Table 1.2 the number of herds, cows and sires and total number of lactations included in the August 2003 genetic evaluation, using a lactation model (*i.e.* based on 305-day yields), are indicated. First lactation 305-day records from 1978 were included in the analyses. The numbers for the test year 2003/2004 are indicated in brackets for the different breeds.

Table 1.2 Number of herds, cows, sires and lactations included in the August 2003 genetic evaluation for the different breeds in South Africa.

Breed	Herds	Registered	Grade	Sires	Lactations
Ayrshire	281	15 163	4 625	708	19 788
	(50)	(1 106)	(157)	(135)	(1 263)
Guernsey	153	4 348	5 073	431	9 421
	(24)	(192)	(161)	(45)	(353)
Holstein	2 848	258 935	188 857	6 149	447 792
	(558)	(14 315)	(9 001)	(1 023)	(23 329)
Jersey	1 707	103 778	33 526	4 213	137 304
	(424)	(11 306)	(2 608)	(848)	(13 914)

(2003/2004 test year)

Using only completed first lactation records in genetic analyses of course has several disadvantages. A cow that did not participate in milk recording during her first lactation or for some reason her first lactation was terminated before 240 days in milk, never receive a breeding value based on performance, regardless of her production in later lactations. Furthermore, the performance of cows in second and third lactations is never included, rendering less accurate genetic evaluations, especially for “late bloomers”. For this reason breeding values of sires may also be biased. The genetic evaluation system for dairy breeds in South Africa was therefore criticized by its participants and the dairy industry. An urgent need for upgrading the genetic methodology as well as for including the estimation of breeding values for somatic cell count to be included in an udder health index, were evident.

1.2 GLOBAL GENETIC EVALUATIONS FOR DAIRY CATTLE

The genetic evaluation of dairy sires and cows for production traits has for many years been based on the analyses of 305-day lactation yields (Lactation Models). The basis of every 305-day yield is a set of test-day yields taken approximately every 35 days in milk (Figure 1.1). Incomplete lactation records are normally extended to a 305-day basis, following a set of well-defined rules.

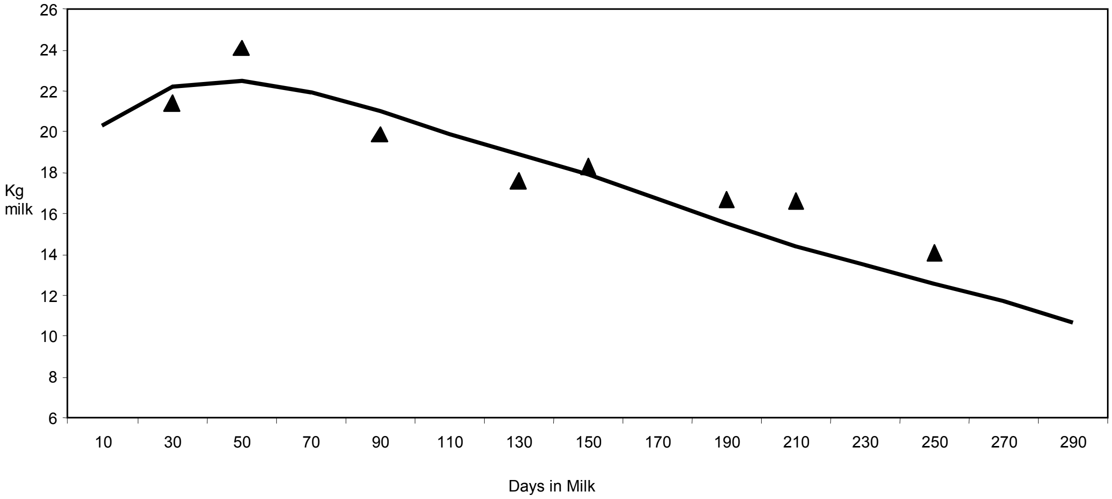


Figure 1.1 A set of test-day yields as basis of a 305-day milk yield.

The accuracy of 305-day measures varies depending on the number of tests that have been combined and the procedures used (Ptak & Schaeffer, 1993). In recent years, interest in test-day records has increased among dairy geneticists and amongst members of the dairy cattle breeding industry (Swalve, 2000), *i.e.* the inclusion of test-day records in genetic evaluations instead of using lactation records, which are calculated from these test-day records (Swalve, 1998).

In Table 1.3 the methodology applied for genetic evaluations of dairy breeds by countries participating in INTERBULL, are listed for production and somatic cell score traits (www-interbull.slu.se (March, 2004)). With regards to production traits, South Africa and Mexico include only first lactation records in their analyses. Four countries include all lactations in genetic evaluations, usually grouping lactations three and following lactations together. The model used by Finland, for example, describe observations of lactations two and later as repeated measures (Lidauer *et al*., 2000). Six countries include the first five parities and the rest (15 countries) include the first three parities in their analyses. Sire Models have been replaced by Animal Models for the evaluation of production traits by all countries. Of the 29 countries for which information was available, nine countries implemented test-day models, of which five uses Random Regression Models, New Zealand uses a Two-step Test-day Model, the DEA (Germany and Austria) uses a Reduced Rank Random Regression Model and Estonia and Switzerland use Fixed Regression Test-day Models. Countries using Test-day Models all use Multi-Lactation Models, meaning that different lactations are included as different traits in the analyses. Twenty countries still use Lactation Models, of which most (15) include the different lactations as repeated measures in single trait analyses. Norway also uses a Repeatability Model, but include all traits in a Multitrait Animal Model. The Czech Republic uses a Single Trait, while Poland uses a Multitrait Multiple Lactation Animal Model, therefore including different lactations as different traits in the analyses.

Table 1.3 Methodology applied for genetic evaluations of the Holstein breed by countries participating in INTERBULL, for production and somatic cell score traits (www-interbull.slu.se (March, 2004)).

Country	**Production**		**Somatic Cell Score**	
	Parities	**Methodology**	**Parities**	**Methodology**
Argentina		Not available		-
Australia	All	ST-RP-AM	3	MT-TD-AM*
Belgium	3	MT-ML-RR-TD-AM	3	ML-RR-TD-AM
Canada	3	MT-ML-RR-TD-AM	3	MT-ML-RR-TD-AM**
Czech Rep	3	ST-ML-AM		-
DEA	All	ST-ML-RRR-TD-AM	3	ST-ML-FR-TD-AM
Denmark	3	ST-RP-AM	3	MT-SM
Estonia	3	ST-ML-FR-TD-AM	3	ST-ML-FR-TD-AM
Finland	10	MT-ML-RR-TD-AM	3	ST-RP-AM
France	3	ST-RP-AM	3	ST-RP-AM
Germany	3	ST-ML-RR-TD-AM	3	ST-ML-RR-TD-AM
Hungary	3	ST-RP-AM	3	ST-RP-AM
Israel	5	ST-RP-AM	5	ST-RP-ML-AM
Italy	3	ST-RP-AM	1	ST-RP-TD-AM
Ireland	5	ST-RP-AM		-
Japan	5	ST-RP-AM	1	ST-FR-TD-AM
Mexico	1	ST-AM		-
Netherlands	3	ST-ML-RR-TD-AM	3	ST-ML-RR-TD-AM
New Zealand	All	ST-RP-TD-AM***		-
Norway (Ays)	3	MT-RP-AM		-
Poland	3	MT-ML-AM		-
Portugal	All	ST-RP-AM		-
Slovenia	5	ST-RP-AM		-
RSA	1	MT-AM		-
Spain	5	ST-RP-AM	5	ST-RP-AM
Sweden	3	ST-RP-AM		-
Switzerland	3	ST-ML-FR-TD-AM	3	ST-ML-FR-TD-AM
Turkey	3	ST-RP-AM		-
UK	5	ST-RP-AM	5	ST-RP-AM
USA	5	ST-RP-AM	5	ST-RP-AM

AM = Animal Model
MT = Multitrait
RRR = Reduced Rank RR
TD = Test-day
FR = Fixed Regression
RP = Repeatability Model
SM = Sire Model
ML = Multi-Lactation
RR = Random Regression
ST = Single Trait

*Multiple Trait Test-day Model
**Somatic cell score included in analysis of production traits
***Two step Test-day Model

For Somatic Cell Score 18 countries participate in INTERBULL, of which Denmark still uses a Sire Model and include Somatic Cell Score in a multitrait analysis with Clinical Mastitis. Ten countries implement Test-day Models, of which Belgium, Canada, Germany and the Netherlands uses Random Regression Models. Australia uses a Multiple Trait Test-day Model, where somatic cell score on different test-days are defined as different traits. Four countries use Fixed Regression Models. For the test-day models, different parities are included as different traits in the analyses, while the seven countries that still implement Lactation Models, include different parities as repeated measures in their genetic evaluations.

1.3 TEST-DAY MODELS

The genetic evaluation of dairy cattle using test-day models has been investigated by several research groups in recent years (Swalve, 1995a, 1998, 2000; Freeman, 1998; Misztal *et al.*, 2000 and Jensen, 2001) and some countries have already implemented routine genetic evaluation of large commercial dairy populations using a Test-day Model (Swalve, 2000; Table 1.3).

Advantages of using Test-day Models compared to the traditional use of yields aggregated over 305 days of lactation (LM), are as follows :

- The most important advantage is probably that a more precise adjustment for temporary environmental effects on the test-day can be made (Ptak & Schaeffer, 1993; Swalve, 1998, 2000; Reents *et al.*, 1998; Mrode *et al.*, 2000 and Jensen, 2001).
- By using test-day yields for genetic evaluation of dairy sires and cows rather than 305-day yields, the problem of extension of test-day yields into a 305-day record, are solved (Ptak & Schaeffer, 1993; Swalve, 2000; Jensen, 2001).
- Projection of records to 305-day yields usually assume a fixed shape of lactation curve for all cows and tend to underestimate 305-day yields from early test-days for more persistent cows and overestimate yields for the less persistent cows. This tends to adversely affect the initial evaluations for young bulls based on part lactation records. Their breeding values are either under or over-predicted relative to breeding values estimated from completed lactations. Test-day models have the advantage that this record-in-progress dip (rip-dip) effect (Jamrozik *et al.*, 1997b) can be avoided (Mrode *et al.*, 2002).
- The accuracy of a cow's genetic evaluation may be improved by using four or more test-day yields per cow per lactation, rather than having only one 305-day measure. With the test-day model cows can be evaluated as long as they have at least one test-day measurement per lactation. Sires are also proven more accurately by having a large number of test-day yields available on their daughters, rather than only one record per daughter, although the increase in accuracy will be due to the better model for test-day yields rather than the number of test yields (Ptak & Schaeffer, 1993).
- Costs of recording dairy cattle performance can be reduced. The supervised recording at monthly intervals is no longer the method of choice in many countries. Rather, mixtures of recording schemes, e.g. pro-longed intervals, a.m. and p.m. alternate recording and owner-sampler recording, have been implemented along with the reference method. The use of test-day records therefore enables records from different schemes to be combined (Swalve, 1998, 2000). Furthermore, milk recording agencies may not need to collect nine or ten test-day yields per cow per lactation for test-day analyses and this could result in lower cost to dairy producers (Ptak & Schaeffer, 1993).
- Specific recording plans will also be allowed making provision for some herds to only contribute milk yields, while in others fat and protein contents are also sampled.
- Compared to only using records of complete lactations, the use of test-day records can reduce the generation interval through frequent genetic evaluations with the latest data. Decreasing the

generation interval has been identified as a main tool to increase genetic progress in dairy cattle breeding (Swalve, 2000).

- With test-day models the shape of the lactation curve can be modelled to differ for each cow (Jamrozik et al., 1997b; Mrode *et al*., 2000 and Bormann *et al*., 2003). In a fixed regression model all cows within a class (e.g. age - parity - season) are assumed to have the same lactation curve. A random regression test-day model allows the shape of the lactation curve to be split into two parts: a general part, accounting for similarities of lactation curves within specified groups (*i.e.* age classes) and an animal specific part (individual deviation from the class average lactation curve). The general part is modelled by a fixed function and the animal part by a random one (Bormann *et al*., 2003).
- Test-day models also allow for the possibility of genetic evaluation for persistency of lactation (Jamrozik *et al*., 1997b; Mrode *et al*., 2000 and Jensen, 2001).
- The effect of pregnancy can be modelled in test-day evaluations, because a test-day record can directly be linked to whether or not a cow is pregnant (Swalve, 2000).
- Using the test-day record database, it should also be possible to provide culling rates on bulls' daughters that were not readily available in the past (Schaeffer, 2000).
- Furthermore, even within a herd-test-day record, further subdivision of herd-test-day into management groups is feasible if they are recorded (Swalve, 2000), for example day of the year, number of days in milk, pregnancy status, medical treatment and number of times milked on the specific test-day. Many of these factors can change for a cow from one test to the next and are difficult to model for 305-day yields (Jamrozik *et al*., 1996).

Disadvantages of using test-day yields include the need to store all of the individual test-day yields on every cow (Ptak & Schaeffer, 1993). Furthermore, the use of test-day observations in the statistical model greatly increases the number of estimated parameters and the amount of data to be analysed, which increases computing time (Ptak & Schaeffer, 1993; Reents *et al*. 1998; Jensen, 2001 and Lidauer *et al*., 2003). Defining the contemporary group as herd–test-day leads to much more levels to be accounted for compared to the herd-year-season classification, commonly used in lactation models (Reents *et al*., 1998). Consequently, the type of test-day model adopted is often dictated by computational considerations (Lidauer *et al*., 2003). Also, tradition has built a strong dependence upon 305-day information, so that 305-day yields would still need to be provided to dairy producers for management purposes (Ptak & Schaeffer, 1993). Schaeffer *et al*. (2000) found it significantly challenging to inform the dairy industry about changes in EBV after changing from a lactation model to a test-day model in Canada.

1.3.1 Types of Test-day Models

Different types of test-day models are used for national dairy cattle evaluations worldwide. A repeatability or fixed regression test-day model was used for German Holstein cattle (Reents & Dopp, 1996; Reents *et al*., 1998) and is currently used for evaluation of production traits in Slovenia (Potocnik *et al*., 2001), Estonia and Switzerland (Table 1.3). Random regression test-day models are used in Canada (Schaeffer *et al*., 2000; Kistemaker, 2003), The Netherlands (De Roos *et al*., 2001), in Germany for the Holstein and Red breeds (Liu *et al*., 2001) and in joined evaluations for Simmental and Brown Swiss in Germany and Austria (Emmerling *et al*., 2002). Finally, a reduced rank random regression model is used in Finland (Lidauer *et al*., 2000), and is being investigated by The Netherlands (De Roos *et al*., 2002) and the UK (Mrode *et al*., 2003). When comparing this literature to the information in Table 1.3, there are some inconsistencies, as the information on the INTERBULL website was probably not yet updated for the relevant countries.

The analysis of test-day records entails the analysis of repeated records on an individual (Jensen, 2001). In general, two approaches may be differentiated in the use of test-day models :

1.3.1.1 Two-Step Test-day Models

In these models the test-day records are corrected for test-day environmental effects, typically using a model with fixed effects only. The residuals from this model are then combined into lactation measures that can be analyzed with traditional models for lactation yields (Jensen, 2001). For countries that have not been storing test-day information for a long time, a switch to a test-day model is facilitated if a two-step model is used, because combining lactation and test-day records is more straightforward than using a direct test-day approach. Combining older lactation data with more recent test-day data in this case is required, because the breeding values of living animals could otherwise change dramatically when switching occurs, primarily because of the effects of older animals that are without test-day records, but are related to living animals with test-day records (Swalve, 2000). These models are routinely used in Australia and New Zealand (Jensen, 2001).

1.3.1.2 One-Step Test-day Models

1.3.1.2.1 Fixed Regression Test-day Models (FRM)

One-step test-day models have been derived from repeatability animal models under which the test-day records within a lactation are included as repeated measurements (Swalve, 2000). The lactation curve is modeled as a fixed effect and the random component of the model is specified as a traditional repeatability model (Jensen, 2001). The stage of lactation is therefore considered only in the fixed effect part of the model, while the random genetic animal effect is modeled as a constant for each day in milk (Lidauer *et al.*, 2003). A simplified scalar version of a test-day model with fixed regressions for records in a single parity could be as follows :

$$y = HTD + \Sigma\, b_i x_i + a + pe + e$$

where HTD = fixed herd – test-day effect
b_i = regression parameters modeling the lactation curve on days in milk
x_i = covariates modeling the lactation curve on days in milk
a = random genetic effect of the animal
pe = permanent environmental effect associated with each cow
e = random residual effect

The regressions are typically nested within classes of fixed effects such as age, season and region. Considering the random effects, the model is a simple repeatability model that assumes constant additive genetic and permanent environmental variances throughout the lactation. It is furthermore assumed that the genetic and permanent environmental correlations between the yield at different days in milk are unity, regardless of the distance between the days chosen (Jensen, 2001). Models like these are generally referred to as Fixed Regression Models (Swalve, 2000).

1.3.1.2.2 Random Regression Test-day Model (RRM)

The random regression test-day model is an extension of the FRM, where the random genetic animal effect can vary for each day in milk and is modeled by a random regression function on days in milk (Lidauer *et al.*, 2003). It is therefore assumed that the shape of the lactation curve is also influenced by random genetic and permanent environmental effects and that the genetic and permanent environmental correlation between yields at different days in milk can be less than one (Jensen, 2001). The chosen random regression function determines the number of equations per animal and thus the computational requirements (Jamrozik *et al.*, 1997c). Because the RRM accounts for individual cow differences in the shape of the curve, it should be more efficient in overcoming or reducing the rip-dip effect compared to a FRM (Mrode *et al.*, 2002). In simple scalar form, the model can be defined as :

$$y = HTD + \Sigma\, b_ix_i + \Sigma\, a_jx_j + \Sigma\, p_kx_k + e$$

where :

HTD = fixed herd – test date effect
b_i = regression parameters modeling the lactation curve on days in milk
x_i = covariates modeling the lactation curve on days in milk
a_j = additive genetic effect corresponding to regression coefficient j
x_j = covariates modeling the additive genetic effect on days in milk
p_k = permanent environmental effect corresponding to regression coefficient k
x_k = covariates modeling the permanent environmental effect on days in milk
e = random residual effect

The different subscripts indicate that the covariates in different parts of the model are not necessarily the same. The covariates x_i can, in principle, be any covariates, but are usually relatively simple functions of days in milk such as polynomials, orthogonal polynomials (e.g. Legendre polynomials) or the parameters of the Wilmink function. RRM can accommodate heterogeneous additive and permanent environmental variances during the lactation. The degree of heterogeneity may depend on the functions chosen to model the trajectory of lactation (Jensen, 2001).

1.3.1.2.3 Reduced Rank Random Regression Test-Day Model (RRRM)

The reduced rank random regression test-day model is a modification of the RRM, where the dimensionality of the test-day model is reduced by using only significant eigenvalues and eigenvectors of the decomposed variance-covariance matrices. They would therefore be computationally less expensive. Lidauer *et al.* (2003) found that changes in ranking of active bulls and cows were small when using a RRRM compared to a RRM and that faster convergence and lower demand of computer memory reduced the computational burden for solving the RRM to a level more similar to the FRM. Druet *et al.* (2003) also commented that models using eigenvectors seemed appealing, because they can reduce the computational difficulty of the model and improve its convergence properties.

1.3.1.2.4 Multiple Trait Test-Day Model

In these models records at each day in milk, or records in classes of days in milk defined in intervals along the lactation, are considered as separate traits. These traits are then analyzed using multivariate methodology (Jensen, 2001). Meyer *et al.* (1989), Pander *et al.* (1992) and Rekaya *et al.* (1995) showed in studies estimating genetic and phenotypic correlations among test-day records applying multiple trait models, 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. The advantage of these models is that no structure is assumed for the (co)variances among records taken at different days in milk. The disadvantage is that many “traits” must be defined, leading to difficulties in estimating the fixed effects and the dispersion parameters pertaining to the additive genetic and permanent environmental effects in the model. Furthermore, because no structure is assumed on the development of (co)variances over time, the parameters can be “jumpy”. This might be due to the relatively small datasets that are typically used for parameter estimation. Such behavior is in contrast to the expectation of a smooth development over the lactation (Jensen, 2001). Compared to a multiple trait test-day model, a RRM estimates variances and covariances smoother and with less bias (Kirkpatrick *et al.*, 1990), needs fewer parameters to describe the same data and provide a method for analyzing independent components of variation that reveal specific patterns of change over time (Huisman *et al.*, 2002).

1.3.1.2.5 Covariance Function Models

A covariance function is a function describing the (co)variance among records/traits that are measured at different days in milk during the lactation (Jensen, 2001). Covariance functions can be equivalent to RRM if the same functions are used (Swalve, 2000). These functions can also be used to model the covariances between different traits such as milk, fat and protein. The advantage of this is that a reduced fit can be used, especially in the multivariate case, and this leads to considerably fewer parameters that must be estimated.

1.3.1.2.6 Character Process Models

These models do not attempt to model the production curve of an animal, but aim at modeling the covariance function (Jensen, 2001 and Meuwissen & Pool, 2001). This contrasts to the RRM, whose primary aim is to model the production curve, and the covariance function results from the estimates of the random regression curves. A limitation of these models is that the curve of breeding values of individuals is not explicitly estimated and may be difficult to obtain (Meuwissen & Pool, 2001).

1.3.1.2.7 Spline Models

A cubic spline is a smooth curve over an interval formed by linked segments of cubic polynomials at certain knot (border) points, such that the whole curve and its first and second differentials are continuous over the interval (Green & Silverman, 1994). Cubic spline models in dairy evaluations consist therefore of a series of cubic polynomials, each defined in an interval on days in milk. They are constructed as such that they are continuous at the knots when moving from one interval to the next. The advantage of splines is that they offer greater flexibility than the functions typically used in RRM and that they do not exhibit the end-effects of a polynomial, which tends to bend more sharply at the extremities (Verbyla *et al*., 1999; Druet *et al*., 2003). Furthermore, they have limited sensitivity to the data and can be relatively easily implemented in mixed model equations (Druet *et al*., 2003). However, they also increase computational demands.

The different test-day models are clearly connected. On the one extreme the FRM assumes a genetic correlation of unity between records obtained at different days in milk. At the other extreme the Multitrait Test-day Model assumes no structure on these covariances, while the Covariance Function Model offers a compromise between these extremes (Jensen, 2001).

Application of test-day models is a formidable task. All test-days are correlated as are all production traits. The choice of the appropriate test-day model to use is not clear and differs across countries (Freeman, 1998). It is unlikely that a single “best model” can be found, since local circumstances might determine what effects should be included in the model. In principle, the model that maximises genetic progress in the population should be chosen for genetic evaluation, even though this may be hard to verify for specific models (Jensen, 2001). If the model is not appropriate, genetic and environmental effects may be confounded (Freeman, 1998).

1.3.2 Comparison of Models for Genetic Evaluation

1.3.2.1 Test-day Models *versus* Lactation Models

1.3.2.1.1 Genetic Parameters

Swalve (1998) summarized from the literature that higher heritabilities are estimated under a two-step approach compared to using LM. For one step approaches, most studies reported smaller heritabilities for test-day models. De Roos *et al.* (2001) explained that the interpretation of 305-day heritabilities is not straightforward, since 305-day yield is not measured on a cow, but a combination of many test-days. Because correlations between different parities and days in milk are highest for the genetic effect, lower for the permanent environmental effect and zero for the residual effect, heritabilities of 305-day traits should be higher than the underlying test-days. According to Serrano *et al*. (2003) repeatability models, either with average lactation somatic cell score or with somatic cell score test-day records, showed lower genetic variances than RRM. High permanent environmental variance relative to the genetic variance, indicate some degree of confounding. This might be due to the lack of pedigree information and insufficient repeated measures per animal.

1.3.2.1.2 Estimated Breeding Values

Already in 1993 Ptak and Schaeffer published correlations of test-day evaluations with 305-day evaluations from 0.87-0.97. Reents *et al.* (1995a) reported that differences between EBVs (Estimated Breeding Values) from FRM and LM diminish with increasing number of daughters. In a study conducted by Reents & Dopp (1996) where EBVs from a multiple lactation FRM was compared to EBVs from a LM where the 1st lactation yield of cows were defined as three 100 day parts and these parts and lactations 2 and 3 were seen as different traits, a correlation of 0.95 was reported for EBVs of first lactation for older bulls with 75 or more daughters. For the younger bulls correlations were lower, especially for bulls with a small number of daughters. Reason for this might be that test-day evaluations for young bulls had much more test-day records incorporated than the respective part lactation EBVs from the LM. Agreement between EBVs for second and third lactations was generally lower than for first lactations. They concluded that the test-day model is much more flexible with respect to incorporation of records in progress or terminated lactations due to culling.

Reenst *et al*. (1998) observed a larger increase in the standard deviations of sires' EBVs (14%) when using a FRM rather than a LM. They found changes in ranking to be substantial, *i.e.* 0.94 for bull EBVs and 0.87 for cow EBVs.

Schaeffer *et al*. (2000) reported correlations of 0.90 (13 Milk Shorthorn sires) to 0.97 (4293 Holstein sires) between milk EBVs of sires from the official LM and RRM in Canada. They concluded that although average changes in EBVs and correlations indicated good agreement between models, the listings of top bulls showed more significant re-rankings. The test-day model, however, is expected to yield more stable EBVs than the LM over time.

Schaeffer *et al*. (2000) and Lidauer *et al*. (2000) showed that moving from a LM to a test-day model has a large impact on EBVs of cows, but less impact on EBVs of bulls. This is because the test-day models allow a better modeling of the herd environment, which improves the accuracy of cows' EBVs. Bulls with a large number of daughters receive accurate EBVs also with the LM. The advantage of test-day models for bulls is therefore more visible in EBVs of bulls with a small number of daughters.

Lidauer *et al*. (2003) found that compared to the LM, EBVs from a RRM gave higher standard deviations : the increase was smallest for active sires (between 2 and 4%), intermediate for young bulls (between 9 and 11%) and highest for young cows (between 15 and 20%). A part of the increase in standard deviations of EBVs for young bulls and young cows can be explained by the better use of information in the test-day models. In this study EBVs from test-day models were compared with EBVs from official national genetic evaluations based on LM. In contrast to other studies that differed in the amount of data, pedigrees and heritabilities (Reenst & Dopp, 1996; Jamrozik *et al*., 1997a,b; Reents *et al*., 1998, Lidauer *et al*., 2000; Schaeffer *et al*., 2000 and Emmerling *et al*., 2002), their study tuned the animals and data carefully to be the same across different models and high correlations were therefore found between EBVs of active bulls from RRM and LM for all traits (0.99 for milk, 0.98 for butterfat and 0.98 for protein, respectively). Overall the correlations reported by Lidauer *et al.* (2003) between active bulls' EBVs obtained from three test-day models (FRM, RRM and RRRM) and the corresponding EBVs from a LM, were higher than reported in the mentioned literature (0.87-0.97). These high correlations suggest almost equal ability of the LM and the test-day models in modeling EBVs of proven bulls. However, for young cows the corresponding correlations were about 0.88, which led to significant re-ranking of cows. Given that the cow EBVs from the test-day model are more accurate, test-day models should have a positive effect on genetic progress due to more accurate selection of bull dams. Correlations were lowest between EBVs from the LM and FRM for young cows (0.84-0.86). They concluded that problems in modeling breeding values for cows with lactations in progress might be the reason for this.

1.3.2.1.3 Genetic Trends

Reents *et al*. (1998) found that the genetic trend is significantly higher from test-day models compared to LM. More young animals were therefore presented in the top-100 bull rankings and also top-1000 cow lists in their study. According to Lidauer *et al*. (2003) RRM and RRRM gave identical genetic trends for milk, butterfat and protein (where heritabilities on a 305-day basis were the same for the different models) and the genetic trends obtained from a LM followed these trends very closely. However, genetic trends from a FRM were slightly different. A part of this discrepancy was due to the model not being able to separate breed differences, partially due to sub-optimal modeling of the fixed effects. This was not a problem with the RRM and RRRM, since lactation curves were estimated for each genetic group, which gave a channel for a breed difference in the shape of the lactation curve.

1.3.2.1.4 Computation

Reents *et al.* (1995b) commented that compared to time requirements for editing and data preparation of test-day records, the difference in CPU time for genetic evaluation was of minor importance. With regards to solving the mixed model equations Lidauer *et al*. (2003) reported that the total computing time was shortest for the LM and increased by a factor of 4.3 when using a FRM and 7.7 when using a RRRM. However, computation for solving the RRM lasted 116 times longer than when solving the LM. The poor convergence when using the RRM, explained the difference in solving time. For this test-day model 2269 iterations were required to reach convergence, whereas for the RRRM, 160 iterations were sufficient. The low rate of convergence in solving the RRM was due to the complexity of the variance-covariance matrices of the animal and permanent environmental effects.

1.3.2.2 Comparison of different test-day models

1.3.2.2.1 Genetic parameters

Swalve (1998) reported that higher heritabilities are estimated under RRM compared to FRM. RRM may have a tendency to underestimate genetic correlation between test-days far apart in days in milk and in different lactations (Van der Werf *et al*., 1998 and Kettunen *et al*., 2000).

1.3.2.2.2 Estimated Breeding Values

Liu *et al.* (1998) compared a FRM with a RRM using Canadian test-day data and found that the RRM had larger variances of EBVs than the FRM, both for cows and bulls. They concluded that the RRM gave a reasonable goodness of fit to test-day data and demonstrated superiority over the FRM in modeling test-day data. Pool & Meuwissen (2001) reported high correlations between EBVs from RRM and FRM with the genetic trend for young bulls being steaper with the FRM (106kg milk/year) compared to the RRM (94kg milk/year). EBVs of young bulls predicted by FRM were at the age of 4.5 years in general lower than at the age of 6 years (*i.e.* EBVs based on records in progress and complete lactations from first *versus* second crop daughters, respectively), while the RRM showed only slightly lower EBVs at the age of six years old. Differences, however, were relatively small.

Reinhardt *et al.* (2002) conducted test-runs using a RRM parallel to the official genetic evaluation system of the time based on FRM for German and Austrian Holstein, Red and Jersey breeds. For young bulls having daughters with incomplete or missing lactations, the RRM resulted in slightly more stable EBVs over time than the FRM, though the difference in stability was marginal. EBVs of both models were highly correlated and the correlations decreased as daughters' lactation information became more complete, indicating different projection of EBVs for missing days in milk and lactations by the two models. Liu *et al*. (2003) confirmed these results with a study comparing EBVs from a RRM with that of a FRM, previously used as the genetic evaluation model for Austrian, German and Luxembourgish breeds for somatic cell score.

According to Mrode *et al*. (2002) PTAs (Predicted Transmitting Abilities) based on initial two test-days compared with those from 10 test-days with both RRM and FRM, resulted in substantial re-ranking in bull and cow proofs. With four to six test-days, good initial predictions of the final PTAs based on 10 test-days were obtained with both models. Generally, the FRM resulted in larger mean under-prediction of cow and bull PTAs with part lactation test-days when compared to RRM. The same trend was observed in terms of mean over-prediction, except for the initial two test-days, where the RRM resulted in higher mean over-predictions, especially in cows with initial high test-day yields, but poor persistency. The use of the initial two test-days in the evaluation of young bulls with a RRM, could therefore still result in a rip-dip effect for some young bulls whose daughters are very persistent or whose initial test-day yields are very high but persistency is very poor.

Lidauer *et al*. (2003) did a close investigation of the young cow group in their study and revealed that the higher standard deviations in EBVs from a FRM was caused by a higher variation in EBVs of cows with less than five test-day records. For all test-day models tested by Lidauer *et al*. in 2003 (a RRM, RRRM and FRM), standard deviations of EBVs were the same for cows without test-day records. Similarly, for cow groups that were eligible to have six test-day records, all of the models gave the same standard deviations of EBVs, which were roughly 25% higher than the standard deviation of EBVs based on pedigree indices. For the FRM, standard deviations of EBVs were inflated for cows with one to four test-day records. These results are, according to Lidauer *et al*. (2003), most likely due to the underlying simplifications in a FRM in that it assumes that heritability and repeatability are constant along the course of lactation.

1.4 CLOSURE

From this literature study it is clear that extensive work has been done to study test-day modeling. Except for comparison of test-day models with 305-day lactation models, several studies have considered the definition of the contemporary group in test-day models, where herd x test-date groupings rendered higher heritability estimates than the traditional herd x year x season groupings (Meyer *et al*., 1989; Ptak & Schaeffer, 1993, Reents *et al*., 1995c and Swalve, 1995b). The shape of the lactation curve fitted (Guo & Swalve, 1995; Rekaya *et al*., 1995; Reinhardt *et al*., 2002; Druet *et al*., 2003; Kistemaker, 2003 and López-Romero & Carabano, 2003) and adjustment for heterogeneous variance (Reents *et al*., 1998; Kistemaker & Schaeffer, 1998; Schaeffer *et al*., 2000; De Roos *et al*., 2001; Lidauer & Mäntysaari, 2001; and Gengler & Wiggans, 2001; 2002) were also discussed thoroughly. Furthermore, many studies obtained genetic parameter estimates for test-day yields (Rekaya *et al*., 1995; Ziu *et al*., 2000; Haile-Mariam *et al*., 2001; Mrode & Swanson, 2001; De Roos *et al*., 2003; Druet *et al*., 2003 and Serrano *et al*., 2003), where smaller heritability at the beginning and end of lactation and correlations ranging from figures over 0.9 for adjacent tests to correlations near 0.7 for distant measurements, were found.

1.5 AIM

The aim of this study was to investigate the suitability of test-day models for genetic evaluation of dairy cattle in South Africa. Test-day records of the first three lactations of Holstein, Jersey, Ayrshire and Guernsey cows participating in the South African Dairy Animal Improvement Scheme were obtained from the INTERGIS (Integrated Registration and Genetic Information System of South Africa). Traits included in the study are milk, butterfat and protein yield/day, as well as somatic cell score as indicator trait for resistance to mastitis.

To reach the aim of the study, the following issues were addressed :

Data Editing : Specifications for test-day records to be included in a national analyses.

Evaluation Model : Should a multiple trait model, including the different lactations as repeated measures and utilizing genetic correlations between traits be implemented *versus* a multiple lactation model, treating different lactations as different traits within every trait ?

Test-Day Model : Should a fixed regression *versus* a random regression test-day model be implemented ?

Genetic Trend Validation : The evaluations will be subjected to three methods for validation of the system developed for national genetic analyses of the dairy breeds in South Africa.

Presentation : In what format should the results be presented to the industry for easy and practical interpretation ?

The results of these analyses will be validated by INTERBULL, as South Africa will participate in INTERBULL test-runs before official release of these breeding values to the dairy industry in South Africa.

1.6 REFERENCES

Bichard, M., 2002. Genetic improvement in dairy cattle – an outsider's perspective. Livest. Prod. Sci. 75, 1-10.

Bormann, J., Wiggans, G.R., Druet, T. & Gengler, N., 2003. Within-herd effects of age at test-day and lactation stage on test-day yields. J. Dairy Sci. 86, 3765-3774.

De Roos, A.P.W., Harbers, A.G.F. & De Jong, G., 2001. Random regression test-day model in The Netherlands. In: INTERBULL Open Meeting, Budapest, 30-31 August/Bull. 27th Int. Bull Eval. Serv., Uppsala, pp. 155-158.

De Roos, A.P.W., Harbers, A.G.F. & De Jong, G., 2003. Genetic parameters for test-day somatic cell score estimated with a random regression model. In: INTERBULL Open Meeting, Rome, 28-30 August/Bull. 31st Int. Bull Eval. Serv., Uppsala, pp. 97-101.

De Roos, A.P.W., Koenen, E.P.C., Harbers, A.G.F. & De Jong, G., 2002. Model Validation and Rank Reduction of covariance matrices in the random regression test-day Model in the Netherlands. In: INTERBULL Open Meeting, Interlaaken, 26-27 May/Bull. 29th Int. Bull Eval. Serv., Uppsala, pp. 91-94.

Druet, T., Jaffrézic, F., Boichard, D. & Ducrocq, V., 2003. Modeling lactation curves and estimation of genetic parameters for first lactation test-day records of French Holstein cows. J. Dairy Sci. 86, 2480-2490.

Emmerling, R., Lidauer, M. & Mäntysaari, E.A., 2002. Multiple lactation random regression test-day model for Simmental and Brown Swiss in Germany and Austria. In: INTERBULL Open Meeting, Interlaaken, 26-27 May/Bull. 29th Int. Bull Eval. Serv., Uppsala, pp. 111-117.

Freeman, A.E., 1998. Dairy Cattle Breeding. Proc. 6th World Congr. Genet. Appl. Livest. Prod., Armidale, New South Wales, Australia. 23, 293-294.

Gengler, N. & Wiggans, G.R., 2001. Heterogeneity in (co)variance structures of test-day yields. In: INTERBULL Open Meeting, Budapest, 30-31 August/Bull. 27th Int. Bull Eval. Serv., Uppsala, pp. 179-183.

Gengler, N. & Wiggans, G.R., 2002. Adjustment for heterogenous genetic and non-genetic (co)variance structures in test-day models using a transformation on random regression effect regressors. In: INTERBULL Open Meeting, Interlaaken, 26-27 May/Bull. 29th Int. Bull Eval. Serv., Uppsala, pp. 79-83.

Green, P.J. & Silverman, B.W., 1994. Non-parametric Regression and Generalised Linear Models. Chapman and Hall, London, England.

Guo, Z. & Swalve, H.H., 1995. Modeling of the lactation curve as a sub-model in the evaluation of test-day records. In: INTERBULL Open Meeting, Prague, 7-8 September/Bull. 11th Int. Bull Eval. Serv., Uppsala, pp. 1-4.

Haile-Mariam, M., Goddard, M.E. & Bowman, P.J., 2001. Estimates of genetic parameters for daily somatic cell count of Australian dairy cattle. J. Dairy Sci. 84, 1255-1264.

Huisman, A.E., Veerkamp, R.F. & Van Arendonk, J.A.M., 2002. Genetic parameters for various random regression models to describe the weight data of pigs. J. Anim. Sci. 80, 575-582.

Jamrozik, J., Schaeffer, L.R. & Dekkers, J.C.M., 1996. Random regression models for production traits in Canadian Holsteins. . In: INTERBULL Open Meeting, Veldhoven, 23-24 June/Bull. 14th Int. Bull Eval. Serv., Uppsala, pp. 124-134.

Jamrozik, J., Kistemaker, G.J., Dekkers, J.C.M. & Schaeffer, L.R., 1997a. Comparison of possible covariates for use in a random regression model for analyses of test-day yields. J. Dairy Sci. 80, 2550-2556.

Jamrozik, J., Schaeffer, L.R. & Dekkers, J.C.M., 1997b. Genetic evaluation of dairy cattle using test-day yield and random regression model. J. Dairy Sci. 80, 1217-1226.

Jamrozik, J., Schaeffer, L.R., Liu, Z. & Jansen, G., 1997c. Mutiple trait random regression test-day model for production traits. In: INTERBULL Open Meeting, Vienna, 28-29 August/Bull. 16th Int. Bull Eval. Serv., Uppsala, pp. 43-47.

Jensen, J., 2001. Genetic evaluation of dairy cattle using test-day models. J. Dairy Sci. 84, 2803-2812.

Kettunen, A.E.A., Mäntysaari, E.A. & Pösö, J., 2000. Estimation of genetic parameters for daily milk yield of primiparous Ayrshire cows by random regression test-day models. Livest. Prod. Sci. 66, 251-261.

Kirkpatrick, M., Lofsvold, D. & Bulmer, M., 1990. Analysis of the inheritance, selection and evolution of growth trajectories. Genetics 124, 979-993.

Kistemaker, G.J. & Schaeffer, L.R., 1998. Adjustment for heterogeneous herd-test-day variances. In: INTERBULL Open Meeting, Rotorua, 18-19 January/Bull. 17st Int. Bull Eval. Serv., Uppsala, pp. 55-59.

Kistemaker, G.J., 2003. The Canadian Test-day Model using Legendre Polynomials. In: INTERBULL Open Meeting, Rome, 28-30 August/Bull. 31st Int. Bull Eval. Serv., Uppsala, pp. 202-204.

Lidauer, M. & Mäntysaari, E.A., 2001. A multiplicative random regression model for test-day data with heterogeneous variance. In: INTERBULL Open Meeting, Budapest, 30-31 August/Bull. 27th Int. Bull Eval. Serv., Uppsala, pp. 167-171.

Lidauer, M., Mäntysaari, E.A. & Strandén, I., 2003. Comparison of test-day models for genetic evaluation of production traits in dairy cattle. Livest. Prod. Sci. 79, 73-86.

Lidauer, M., Mäntysaari, E.A., Strandén, I. & Pösö, J., 2000. Multiple-trait random regression test-day model for all lactations. In: INTERBULL Open Meeting, Bled, 14-15 May/Bull. 25th Int. Bull Eval. Serv., Uppsala, pp. 81-86.

Liu, Z., Jamrozik, J. & Jansen, G., 1998. A comparison of fixed and random regression models applied to dairy test-day production data. In: INTERBULL Open Meeting, Rotorua, 18-19 January/Bull. 17th Int. Bull Eval. Serv., Uppsala, pp. 60-63.

Liu, Z., Reinhardt, F., Bünger, A., Dopp, L. & Reents, R., 2001. Application of a random regression model to genetic evaluations of test-day yields and somatic cell scores in dairy cattle. In: INTERBULL Open Meeting, Budapest, 30-31 August/Bull. 27th Int. Bull Eval. Serv., Uppsala, pp. 159-166.

Liu, Z., Reinhardt, F., Bünger, A., Jaitner, J. & Reents, R., 2003. Genetic evaluation of somatic cell scores using a random regression test-day model for a very large dairy cattle population. In: INTERBULL Open Meeting, Rome, 28-30 August/Bull. 31st Int. Bull Eval. Serv., Uppsala, pp. 88-91.

Liu, Z., Reinhardt, F. & Reents, R., 2000. Parameter estimates of a random regression test-day model for first three lactation somatic cell scores. In: INTERBULL Open Meeting, Verden, 22-23 October/Bull. 26th Int. Bull Eval. Serv., Uppsala, pp. 61-65.

López-Romero, P. & Carabano, M.J., 2003. Comparing alternative random regression models to analyse first lactation daily milk yield data in Holstein-Friesian cattle. Livest. Prod. Sci. 82, 81-96.

Loubser, L.F.B, Banga, C.B., Scholtz, M.M. & Hallowell, G.J. (Eds.), 2001. Dairy Herd Improvement in South Africa. The use of breeding values in the South African Dairy Industry. 25.

Meyer, K., Graser, H.-U & Hammond, K., 1989. Estimates of genetic parameters for first lactation test-day production of Australian Black and White cows. Livest. Prod. Sci. 21, 177-199.

Meuwissen, T.H.E. & Pool, H., 2001. Auto-regressive versus random regression test-day models for the prediction of milk yield. In: INTERBULL Open Meeting, Budapest, 30-31 August/Bull. 27th Int. Bull Eval. Serv., Uppsala, pp. 172-178.

Misztal, I., Strabel, T., Jamrozik, J., Mäntysaari, E.A. & Meuwissen, T.H.E., 2000. Strategies for estimating parameters needed for different test-day models. J. Dairy Sci. 83, 1125-1134.

Moore, S.J., 2004. South African National Dairy Animal Improvement Scheme : Annual Report 2004. Vol 24.

Mrode, R.A. & Swanson, G.J.T., 2001. Genetic parameters for test-day somatic cell counts for the first three lactations using a random regression model. In: INTERBULL Open Meeting, Budapest, 30-31 August/Bull. 27th Int. Bull Eval. Serv., Uppsala, pp. 193-196.

Mrode, R.A., Swanson, G.J.T. & Lindberg, C.M., 2000. Efficiency of test-day models in genetic evaluation with part lactation information. In: INTERBULL Open Meeting, Bled, 14-15 May/Bull. 25th Int. Bull Eval. Serv., Uppsala, pp. 87-89.

Mrode, R.A., Swanson, G.J.T. & Lindberg, C.M., 2002. Efficiency of part lactation test-day records for genetic evaluations using fixed and random regression models. Anim. Sci. 74, 189-197.

Mrode, R.A., Swanson, G.J.T. & Paget, M.F., 2003. Implementation of a test-day model for production traits in the UK. In: INTERBULL Open Meeting, Rome, 28-30 August/Bull. 31st Int. Bull Eval. Serv., Uppsala, pp. 193-196.

Pander, B.L., Hill, W.G. & Thompson, R., 1992. Genetic parameters of test-day records of British Holstein-Friesian heifers. Anim. Prod. 55, 11-21.

Pool, M.H. & Meuwissen, T.H.E., 2001. Effects of random regression test-day models on EBVs and genetic trends in persistency. In: INTERBULL Open Meeting, Budapest, 30-31 August/Bull. 27th Int. Bull Eval. Serv., Uppsala, pp. 184-188.

Potocnik, K., Krsnik, J., Stepec, M. & Dolinar, A., 2001. Developments in prediction of breeding values in Slovenia. In: INTERBULL Open Meeting, Budapest, 30-31 August/Bull. 27th Int. Bull Eval. Serv., Uppsala, pp. 107-111.

Ptak, E. & Schaeffer, L.R., 1993. Use of test-day yields for genetic evaluation of dairy sires and cows. Livest. Prod. Sci. 34, 23-34.

Reents, R., Dekkers, J.C.M. & Schaeffer, L.R., 1995a. Genetic evaluation with a mutiple lactation test-day model for SCS and production traits. In: INTERBULL Open Meeting, Prague, 7-8 September/Bull. 11th Int. Bull Eval. Serv., Uppsala, pp. 1-6.

Reents, R., Dekkers, J.C.M. & Schaeffer, L.R., 1995b. Genetic evaluation for somatic cell score with a test-day model for multiple lactations. J. Dairy Sci. 78, 2858-2870.

Reenst, R., Jamrozik, J., Schaeffer, L.R. & Dekkers, J.C.M., 1995c. Estimation of genetic parameters of somatic cell score. J. Dairy Sci. 78, 2847-2857.

Reents, R. & Dopp, L., 1996. Genetic evaluation for dairy production traits with a test-day model for multiple lactations. In: INTERBULL Open Meeting, Veldhoven, 23-24 June/Bull. 14th Int. Bull Eval. Serv., Uppsala, pp. 113-117.

Reents, R., Dopp, L., Schmutz, M. & Reinhardt, F., 1998. Impact of application of a test-day model to dairy production traits on genetic evaluations of cows. In: INTERBULL Open Meeting, Rotorua, 18-19 January/Bull. 17th Int. Bull Eval. Serv., Uppsala, pp. 49-54.

Reinhardt, F., Liu, Z., Bünger, A., Dopp, L. & Reents, R., 2002. Impact of Application of a Random Regression Test-day Model to Production Trait Genetic Evaluations in Dairy Cattle. INTERBULL Open Meeting, Interlaaken, 26-27 May/Bull. 29th Int. Bull Eval. Serv., Uppsala, pp. 103-107.

Rekaya, R., Béjar, F., Carabano, M.J. & Alenda, R., 1995. INTERBULL Open Meeting, Prague, 7-8 September/Bull. 11th Int. Bull Eval. Serv., Uppsala, pp. 1-8.

Schaeffer, L.R., Jamrozik, J., Kistemaker, G.J., Van Doormaal, B.J., 2000. Experience with a test-day model. J. Dairy Sci. 83, 1135-1144.

Serrano, M., Pérez-Guzmán, M.D., Montoro, V. & Jurado, J.J., 2003. Genetic analysis of somatic cell count and milk traits in Manchega ewes. Mean lactation and test-day approaches. Livest. Prod. Sci. 84, 1-10.

Swalve, H.H., 1995a. Test-day models in the analysis of dairy production data – a review. Arch. Tierz. 38, 591-612.

Swalve, H.H., 1995b. The effect of test-day models on the estimation of parameters and breeding values for dairy yield traits. J. Dairy Sci. 78, 929-938.

Swalve, H.H., 1998. Use of test-day records in genetic evaluation. Proc. 6th World Congr. Genet. Appl. Livest. Prod., Armidale, New South Wales, Australia. 23, 295-302.

Swalve, H.H., 2000. Symposium: Test-Day Models. Theoretical Basis and Computational Methods for Different Test-Day Genetic Evaluation Methods. J. Dairy Sci. 83, 1115 – 1124.

Van der Werf, J.H.J., Goddard, M.E. & Meyer, K., 1998. The use of covariance functions and random regressions for genetic evaluation of milk production based on test-day records. J. Dairy Sci. 81, 3300-3308.

Van der Westhuizen, J., 2002. South African National Dairy Animal Improvement Scheme. Annual Report 2002. Vol 22, 4.

Verbyla, A.P., Cullis, B.R., Kenward, M.G. & Welham, S.J., 1999. The analysis of designed experiments and longitudinal data using smoothing splines. J. R. Stat. Soc. Ser. C Appl. Stat. 48, 269-300.

CHAPTER 2

CRITICAL ASSESSMENT AND EVALUATION

2.1 INTRODUCTION

This study was undertaken to investigate the suitability of test-day models for national genetic evaluations of dairy cattle in South Africa, using the available data and software resources. As indicated in Chapter 1 the genetic evaluation system for dairy breeds in South Africa was criticized by its participants and the dairy industry and an urgent need for upgrading the genetic methodology, as well as inclusion of estimation of breeding values for somatic cell count in an udder health index, were required.

The strategy for this study was based on five main aspects, as specified in Chapter 1, including the editing and specifications of the data for national genetic analyses, the evaluation of the most appropriate methodology, model specification, genetic trend validations and presentation of the results to the dairy industry. The research output of this study was compiled into five articles, which is presented as part of this thesis, as well as three short papers, presented at national congresses in the duration of the project.

The aim of this chapter is therefore to discuss the rationale and strategy followed during the research to address the aspects as indicated above and provide and evaluate the final outcome of the study.

2.2 LIMITATIONS AND CHALLENGES

Test-day genetic evaluations on a national basis are expensive in terms of computer memory and time. Available soft- and hardware dictated to a large extent the type of test-day models implemented and the lactation curves fitted in this study.

At the initiation of this study, the most technological advanced computer available in the Quantitative Genetics Division of the Agricultural Research Council (ARC) for national genetic evaluations, was a dual processor computer with two GHz Zion processors and 6G RAM. Available hardware was therefore adequate for implementation of test-day models, specifically fixed regression test-day models.

The software available for prediction of breeding values on a national scale using random regression test-day models, was not available in or affordable by South Africa. Random regression test-day models are currently the “Rolls Royce” models for genetic evaluation of production traits in dairy breeds and would have been the best methodology available to use. This methodology allows for the genetic ability of a cow (as well as permanent environmental effects) to change over the trajectory of the lactation by modeling these effects with random regressions. Instead of predicting a single production breeding value for the animal, a curve is therefore predicted and this also allows genetic evaluations for persistency of lactation. Several software packages are available to the ARC for variance component estimation (ABTK (Golden *et al*., 1992), ASREML (Gilmour *et al*., 2002), VCE4 (Groeneveld (1997), VCE5 (Kovač & Groeneveld, 2003)), however no software was available in South Africa for the prediction of breeding values based on random regression test-day models on a national scale. Thus, random regression test-day models were not yet an option for national genetic evaluations of dairy breeds in South Africa. The software package PEST (Groeneveld, 1990) was, however, available for prediction of breeding values and fixed regression test-day models could easily be modeled with PEST. PEST was therefore used for the prediction of breeding values, based on fixed regression test-day models, where the lactation curve of cows belonging to the same subclass could be modeled by fixed regressions, for dairy breeds in South Africa.

Since 2001 the South African Dairy Animal Improvement Scheme made provision for herds to only contribute milk yields to the Scheme (Personal Communication, Dr J Van der Westhuizen). It was therefore imperative to implement a multitrait model including milk, butterfat and protein yields in the same evaluation, to be able to predict breeding values for missing butterfat and protein yields, based on the genetic correlations of these traits with milk yield. It was furthermore long overdue for genetic evaluations to be based on more than only the first lactation, as was the case with estimation of BLUP breeding values since 1987 in South Africa. It was therefore decided that test-day records of the first three lactations would be included in genetic evaluations in future. As correlations between lactations are not in unity, lactations should be treated as different traits. This study therefore commenced with treating different lactations as different traits, rendering 3x3 covariance matrices for somatic cell score and 9x9 covariance matrices for the production traits to be estimated. However, convergence could not be accomplished for the production traits and hence it was decided to include test-day records within and across lactations as repeated measures of the trait and include milk, butterfat and protein yields in a multitrait evaluation, thus doing Fixed Regression Multiple Lactation Multitrait Test-Day BLUP Animal Models for genetic evaluations of South-African dairy cattle.

2.3 MODEL INFORMATION

An essential step in model design is the definition of the contemporary group and the stipulation of fixed and random effects to be included in the model. Hence the following discussion :

2.3.1 Contemporary Group

The contemporary group provides a way to compare animals fairly. The main advantage of test-day models is the direct correction for fixed effects and especially of fixed effects whose impact changes over time, *i.e.* over several test-days within a lactation of a cow. In all kinds of models that evaluate dairy performance records, a herd – year – season (HYS) effect is commonly used to account for the effects of the individual herd, the year and the season of calving and the interactions amongst them. In test-day models HYS is replaced by the herd – test date (HTD) effect. The HTD effect obviously accounts for the effects of the herd and the year of production and covers also the season of production. The effect of parity is also needed in models for data from multiple lactations (Swalve, 2000). For the test-day model implemented in this study, the contemporary group was therefore defined as herd - test date - parity - number of milkings. Cows that were subjected to exactly the same environmental conditions, were therefore grouped together.

2.3.2 Age of the cow at calving

Age at calving accounts for a large part of the variation in daily milk, butterfat and protein yields. Cows that calved at younger ages in lactations one and two produce less milk compared to cows calving at older ages, over the entire lactation. Second and third parity cows also have higher test-day yields than heifers, for most stages of the lactation (Mostert *et al.*, 2001). Age at calving was therefore included as a fixed effect in the model to compensate for these effects. Age classes were nested within parity because high variation of age existed within parities, causing the age distributions of cows of different parities to overlap.

2.3.3 Calving Interval

As shorter calving intervals have a larger depression on daily production compared to longer calving intervals (*i.e.* smaller fetus) on a given day in milk (Reents & Dopp, 1996), it was important to model the effect of calving interval on yield to ensure fair comparisons amongst cows. Calving interval classes were allocated using standard deviation units, as follows :

Code 1 : All records of Lactation 1.

Code 2 : All records of Lactation 2 with unknown calving intervals (no Lactation 1 records available).

Code 3 : All records of Lactation 2 where the calving interval between Lactations 1 and 2 was less than -1 standard deviation unit from the average calving interval.

Code 4 : All records of Lactation 2 where the calving interval between Lactations 1 and 2 was within 1 standard deviation unit from the average calving interval.

Code 5 : All records of Lactation 2 where the calving interval between Lactations 1 and 2 was greater than 1 standard deviation unit from the average calving interval.

Code 6 : All records of Lactation 3 with unknown calving intervals (no Lactation 2 records available).

Code 7 : All records of Lactation 3 where the calving interval between Lactations 2 and 3 was less than -1 standard deviation unit from the average calving interval.

Code 8 : All records of Lactation 3 where the calving interval between Lactations 2 and 3 was within 1 standard deviation unit from the average calving interval.

Code 9 : All records of Lactation 3 where the calving interval between Lactations 2 and 3 was greater than 1 standard deviation unit from the average calving interval.

2.3.4 Regression Function

The stage of lactation is a key consideration of any test-day model, because it is well-known that the relationship between the stage of lactation and production is curvilinear (Swalve, 2000). The lactation curve should therefore be modeled in test-day evaluations. Usually relatively simple functions on days in milk are used, such as polynomials, orthogonal polynomials (e.g. Legendre polynomials) or models based on biological knowledge of the milk secretion process, e.g. the Wilmink Function (Wilmink, 1987) or Woods Function (1967). The regression is typically nested within classes of fixed effects such as age and season (Jensen, 2001). Because the Wilmink curve could easily be modeled on days in milk with PEST, this was the regression function chosen for the study. The curve was nested within lactation and season, so that six curves (3 lactations x 2 seasons) were fitted for each breed.

The Wilmink Function is defined as follows :

$$y_t = \beta_0 + \beta_1 t + \beta_2 \exp(-0.05\ t)$$

Where

y_t = milk, butterfat or protein yield or somatic cell score on fixed day t

β_0, β_1 and β_2 = coefficients of the lactation curve function

2.3.5 Permanent Environment

It is essential to fit a random permanent environmental effect associated with each cow in test-day models due to a common effect of the environment associated with all test-day records of a cow in a lactation. In fixed regression test-day models constant additive genetic and permanent environmental variances throughout the lactation are assumed (Jensen, 2001). Initially the permanent environmental effect was fitted within lactation, which produced beautiful heritabilities, but concern was expressed by INTERBULL about low correlations from South Africa with countries sharing a reasonable number of common sires, especially with regards to protein yield. As the permanent environmental effect accounts for environmental similarities between test-day records of the same cow, ranking of sires might be influenced when permanent environmental variance is fitted across lactations opposed to only within lactation, because different sires are linked to different numbers of test-day records per daughter per lactation. The permanent environmental effect was then fitted across lactations, which caused a drastic decline in the additive genetic variance, but still yielded reasonable heritability estimates in comparison with other literature estimates based on fixed regression test-day models. The problem of low correlations between South Africa and other countries was solved as well.

2.3.6 Calving Year

It was found that the contemporary group effect did not adequately compensate for the effect of calving year on production or somatic cell score. Calving year was therefore added to the model at a later stage to account for the phenotypic trend over years on production and somatic cell score.

2.4 ADJUSTING FOR HETEROGENEOUS VARIANCES

The Fixed Regression Test-Day Model assumes equal variances at all days in milk. However, tests at the beginning and at the end of lactations have higher variances than tests in the middle of lactation. Furthermore, first lactations have lower mean and variances compared to second and third lactations. A modification was therefore implemented in the model to reduce the effect of deviating from this assumption. This was investigated for the Jersey breed and recommended to be done for the other breeds as well.

2.5 PRESENTATION TO THE INDUSTRY

For the Jersey breed breeding values based on test-day records of the first three lactations were introduced to the industry for the first time early in 2004 and for the Holstein, Ayrshire and Guernsey breeds at the end of 2004. To avoid confusion it was decided to still present breeders with 305-day production breeding values (kg/lactation), and not test-day breeding values (kg/day) which are now the product of the genetic evaluation. Care was taken to inform herd societies, semen agents and breeders of the changes in the genetic evaluation of dairy breeds *via* letters, popular publications, personal communications and presentations. As can be imagined changing from a Lactation Model, using only 305-day first lactation records to a Fixed Regression Test-Day Model using test-day records of the first three lactations and also including short lactations that were previously discarded, led to significant re-ranking of animals which had to be explained to an industry that doesn't welcome change. This was perhaps the main challenge of the study! Passing the INTERBULL trend validation tests, however, was a huge achievement and assisted in receiving acceptance from the Industry.

2.6 CRITICAL ASSESSMENT

The Fixed Regression Multiple Lactation Multitrait Test-Day BLUP Animal Model implemented for genetic evaluations of South-African dairy cattle, is quite a simple model, estimating breeding values from the average of yields within and across lactations. This Model is based on two assumptions, which are well-known to be incorrect :

- Heritability and repeatability are constant along the course of lactation.
- Genetic correlations are unity amongst lactations.

Adjusting for heterogeneous variances due to days in milk and parity alleviate the deviation from these assumptions. Still, it would have been more effective to treat first lactation yields as a different trait from second and third lactation yields, as the genetic correlations of first lactation with second and third lactations are rather low. The genetic correlation between second and third lactations is however, near unity. These yields could therefore be grouped together, rendering 6x6 covariance matrices to be estimated for the production traits which might have reached convergence.

This Model is also subjected to RIP-DIP effects when breeding values are based only on yields early in the lactation. Breeding values of young sires and their daughters will therefore be biased under certain circumstances, eg. where early yields are high, but persistency is low (overestimated), or where early yields are low, but persistency is high (underestimated). Mrode *et al.* (2002), however, showed that Random Regression Models will also estimate biased breeding values due to the RIP-DIP effect.

Modeling the effect of days-carried-calf instead of calving interval would have been more efficient, if insemination dates or expected calving dates were available. The effect of pregnancy on test-day production of two cows with the same calving interval, but different expected calving dates, will differ, as pregnancy status affects how much of the nutrient intake of a cow is allocated for production. Furthermore, the effect of pregnancy on production when the fetus is for example, 90 days old, is expected to be similar for two cows that have different calving intervals.

2.7 CONCLUSIONS

A Fixed Regression Multiple Lactation Multitrait Test-Day BLUP Animal Model has successfully been implemented for genetic evaluation of dairy breeds in South Africa. Genetic evaluation is now based on yields in the first three lactations, which is a considerably more precise measure of lifetime production compared to only first lactation yields, as was previously used. Breeding values for somatic cell score are also now available to the industry for inclusion in udder health indices. Trend validation tests were successful for all traits and breeds except for somatic cell score of the Guernsey breed, due to insufficient data for this trait. This methodology was therefore accepted by INTERBULL, allowing South Africa to participate in international genetic evaluations. This enables the ARC to supply the dairy industry with MACE breeding values, which is invaluable knowledge for the importation of semen, embryos and foreign animals, as it indicates how foreign sires will perform in South Africa, even without having any daughters in South Africa yet. As the major dairy countries recently changed to test-day models, South Africa's genetic methodology is now more comparable to that of the leading dairy countries of the world.

2.8 REFERENCES

Golden, B. L., Snelling, W. M. & Mallinckrodt, C. H., 1992. Animal Breeder's Tool Kit User's Guide. Colorado State Univ. Agric. Exp. Sta. Tech. Bull. LTB92-2.

Gilmour, A.R., Gogel, B.J., Cullis, B.R., Welham, S.J. & Thompson, R., 2002. ASReml User Guide Release 1.0. VSN International Ltd, UK.

Groeneveld, E., 1990. PEST User's Manual. Institute for Animal Science, FAL, D31535 Neustadt, Germany.

Groeneveld, E., 1997. VCE4 User's Guide and Reference Manual. Institute for Animal Science, FAL, D31535 Neustadt, Germany.

Jensen, J., 2001. Genetic evaluation of dairy cattle using test-day models. J.Dairy Sci. 84, 2803-2812.

Kovač, M. & Groeneveld, E., 2003. VCE-5 User's Guide and Reference Manual. Version 5.1. University of Ljubljana, Slovenia & Institute for Animal Science, FAL, D31535 Neustadt, Germany.

Mostert, B.E., Theron, H.E. & Kanfer, F.H.J., 2001. The effect of calving season and age at calving on production traits of South African dairy cattle. S. Afr. J. Anim. Sci. 31, 205-214.

Mrode, R.A., Swanson, G.J.T. & Lindberg, C.M., 2002. Efficiency of part lactation test-day records for genetic evaluations using fixed and random regression models. Anim. Sci. 74, 189-197.

Reents, R. & Dopp, L., 1996. Genetic evaluation for production traits with a test-day model for multiple lactations. Proc. Interbull Meeting, Veldhoven, The Netherlands, June 23-24.

Swalve, H.H., 2000. Symposium: Test-day Models. Theoretical basis and computational methods for different test-day genetic evaluation methods. J. Dairy Sci. 83, 1115-1124.

Wilmink, J.B.M., 1987. Adjustment of test-day milk, fat and protein yield for age season and stage of lactation. Livest. Prod. Sci. 16, 335-348.

CHAPTER 3

THE USE OF FIXED REGRESSION TEST-DAY MODELS IN THE ESTIMATION OF VARIANCE COMPONENTS FOR SOMATIC CELL SCORE OF SOUTH AFRICAN HOLSTEIN AND JERSEY CATTLE

B.E. Mostert#, C.B. Banga, H.E. Theron, E. Groeneveld[1] & F.H.J. Kanfer[2]
ARC-LBU, Private Bag X2, Irene, South Africa
[1]Institute for Animal Husbandry and Animal Behaviour, FAL, Mariensee, Germany
[2]Department of Statistics, University of Pretoria, Pretoria, 0002, South Africa

---

3.1 ABSTRACT

Two fixed regression test day models were applied for variance component estimation and prediction of breeding values for somatic cell score, using test-day records of the first three lactations of South African Holstein and Jersey cows. The first model (ML-model) considered the test-days of the different lactations as different traits in a multiple trait animal model and the second analysis (RM-model) treated later lactation records as repeated measures of the first lactation. Heritabilities from the ML-model were low for both breeds – Holsteins : 0.050, 0.070 and 0.069 for lactations 1, 2 and 3 respectively, and Jerseys : 0.039, 0.044 and 0.046 for lactations 1, 2 and 3, respectively. Lactations 2 and 3 were highly correlated (0.97 for Holsteins and 1.00 for Jerseys), with lower correlations between lactations 1 and 2 (0.83 for Holsteins and 0.81 for Jerseys) and the lowest correlations between lactations 1 and 3 (0.79 for Holsteins and 0.75 for Jerseys). The RM-model estimated heritabilities of 0.19 for Holsteins and 0.18 for Jersey somatic cell scores. Rank correlations between breeding values from the ML-model and RM-model of proven sires, indicated that minor changes in the rank occur between prediction of breeding values from the two models. Although genetic correlations between parities are not unity, the RM-model estimates more competetive variances and requires extensively less computer time to predict breeding values.

Keywords : breeding values, INTERBULL, MACE, mastitis, Wilmink curve

#Corresponding author : bernice@arc.agric.za

3.2 INTRODUCTION

Mastitis is widely known as the most economically devastating disease in dairy cattle. It causes severe financial losses, mainly due to reduced milk yield and quality, treatment expenses and replacement costs. Improvements in management *per se* cannot control mastitis, as environmental organisms causing the disease cannot be eradicated. In addition, economic losses due to mastitis may increase due to the unfavourable genetic correlation between milk yield and mastitis (Emanuelson *et al*., 1988; Simianer *et al.,* 1991; Mrode and Swanson, 1996). Resistance to mastitis is therefore a major goal in dairy cattle improvement.

Direct selection for this trait is difficult to carry out because, as is the case with other health traits, occurrence of mastitis is not routinely recorded in most dairy recording schemes and it is lowly heritable. Somatic cell counts in milk (SCC) are widely used as a means of managing udder health and are routinely recorded in most dairy recording schemes. A high positive genetic correlation has been found to exist between SCC and incidence of mastitis (Emanuelson *et al.,* 1988; Rogers *et al*., 1998; Nash *et al*., 2000; Boichard & Rupp, 2001; Kadarmideen & Pryce, 2001). SCC can therefore justifiably be used as an indirect means of measuring and improving resistance to mastitis through breeding. Consequently, genetic evaluation for SCC is now performed in many countries and international genetic evaluations for SCC are already in place.

Somatic cell count testing was initiated in the South African Milk Recording Scheme in 1995 and currently more than 80% of the milk-recorded herds are routinely tested for SCC. Since 1996, most of the SCC data has been stored on the Integrated Registration and Genetic Information System of South Africa (INTERGIS). The aim of this study was to develop methodology and to estimate (co)variance matrices for the estimation of breeding values for somatic cell count on a national basis for South African Holstein and Jersey dairy cattle.

3.3 MATERIALS AND METHODS

Records from multiple lactations can be considered as different traits in a multiple-trait animal model or by treating later lactation records as repeated measures of the first lactation, assuming a genetic correlation of 1 across lactations (Reents *et al*., 1995a). Data consisted therefore of test-day records of somatic cell count for lactations 1 - 3 of Holstein and Jersey cows calving from 1995 to 2002. These records were obtained from the INTERGIS for cows participating in the South African Dairy Animal Improvement Scheme. The number of test-day records available was 2 385 207 for the Holsteins and 718 016 for the Jerseys. Basic edits included deletion of records with unknown herds, unknown birth dates, test-days recorded before six days in milk or after 305 days in milk, records of crossbred cows and age restrictions within lactations to ensure reasonable calving ages in a specific lactation (20 – 42 months for lactation 1, 30 – 54 months for lactation 2 and 40 – 66 months for lactation 3 for Holsteins and 17 – 40 months for lactation 1, 29 – 53 months for lactation 2 and 41 – 67 months for lactation 3 for Jerseys). Records deleted amounted to 7% of the Jersey and 21% of the Holstein databases. The remaining records were used to ultimately estimate BLUP breeding values for somatic cell score.

In order to ensure a well-linked data structure for variance component estimation, the following data selections were carried out :

Only records with both parents known were used. It was assumed that records of somatic cell counts of 0 were not measured and were therefore deleted. Contemporary groups were then defined as Herd x Test-Date x Parity x Number of Milkings/day. 33% of these contemporary groups of the first parity were randomly selected for the Holstein and 60% for the Jersey analyses. The remaining records were subjected to the following requirements :

- Cows included in the analysis should have a first parity
- There should be daughters of at least two sires in a contemporary group
- The contemporary group should consist of at least five records
- Each sire should be represented in at least three contemporary groups
- Each lactation was divided into 10 stages of 30 days each, except for the 10th stage which consisted of 35 days (270-305). Only one test-day record per stage per cow was included in the analyses.
- A cow should have at least three test-days per lactation to be selected.
- A sire should have at least six daughters in the first lactation.

This selection was based on the study of Haile-Mariam *et al.* (2000).

Data available for variance component estimation then amounted to 111 540 test-day records of 32 075 Holstein cows for the first parity, 115 907 records of 17 441 Holstein cows for the second parity and 48 170 records of 7 534 Holstein cows for the third parity. This amounts to an average of 3.5 records/cow in the first parity, 6.6 in the second parity and 6.4 in the third parity of the Holstein breed, representing 728 sires, 28 558 dams and 539 herds. For the Jerseys the selection ended up with 87 381 test-day records of 17 758 cows in the first parity, 68 845 records of 10 054 cows in the second parity and 31 203 records of 4 650 cows in the third parity. This amounts to an average of 4.9 test-day records/cow in the first parity, 6.8 in the second and 6.7 in the third parity, representing 473 sires, 14 941 dams and 261 herds. Somatic Cell Count (1000's/ml) (SCC) was transformed to log(SCC) for each test-day to achieve normality. This will be referred to as Somatic Cell Score (SCS). (Co)variance components were estimated using VCE4 (Groeneveld & Garcia-Cortes, 1998) with the following genetic model :

$$y_{ijklmn} = \mu + HTDLM_{im} + A_{jm} + PE_{jm} + S_{km} + AC_{lm} + wilmink(S_{km}) + e_{ijklmn}$$

Where

y_{ijklmn} = n^{th} test-day SCS of cow j in lactation m,
μ = mean yield
$HTDLM_{im}$ = fixed effect of herd x test-date x parity x number of milkings group
A_{jm} = animal additive genetic effect
PE_{jm} = permanent environmental (random) effect within lactation to account for common effects of environment associated with all test-day records of cow j in lactation m
S_{km} = fixed effect of calving season in lactation m
AC_{lm} = fixed effect of age class in lactation m
$wilmink(S_{km})$ = Wilmink curve (Wilmink, 1987) modeled on days in milk within season in lactation m (regression)
e_{ijklmn} = random residual error

Season was defined as winter (April – September) *versus* summer (October – March), while the same age classes were allocated as in the derivation of standard lactation curves by Mostert *et al.* (2001).

Two models were applied for analysis of these test-day records. The first model (ML-model) considered the test days of the different lactations as different traits in a multiple-trait animal model and the second analysis (RM-model) treated later lactation records as repeated measures of the first lactation, fitting the permanent environmental effect, as well as the Wilmink curve within lactation. Pedigrees were traced back for four generations, ending with 90 468 animals in the pedigree for the Holstein and 45 320 animals in the pedigree for the Jersey analyses.

BLUP breeding values were estimated using PEST (Groeneveld & Kovac, 1990) by applying pedigrees that included genetic groups, allocated according to year of birth, selection intensity and country of birth. A breeding value index for the ML-model was estimated as the average of the three lactations (MLeq), as well as weighing the lactations on the basis of the Canadian genetic evaluation system, *i.e.* 0.25 for lactation 1, 0.65 for lactation 2 and 0.10 for lactation 3 (MLweig). Rank correlations were done for sires having at least 50 daughters in 10 herds (proven sires), as well as for sires with less reliable breeding values (less than 50 daughters in 10 herds), using SAS (1996).

3.4 RESULTS AND DISCUSSION

A test-day model can be defined as a statistical procedure which considers all genetic and environmental effects directly on a test-day basis (Reents & Dopp, 1996). The most important systematic effects of the environment on SCS that should be removed during genetic evaluations, have been identified as parity, stage of lactation, age of cow, herd and month of calving (Boettcher *et al.*, 1992; Emanuelson & Perrson, 1984; Harmon, 1994; Kennedy *et al.*, 1982 and Schutz *et al.*, 1990). Short-term environmental effects, such as age of the sample, calibration of measurement equipment, infection pressures, improper working of milking equipment and milking practices which influences SCS (Kennedy *et al.*, 1982; Sethar *et al.*, 1979), can partially be accounted for by using test-day models, grouping contemporaries by herd-test date instead of classification by herd-year-season of calving (Reents *et al*., 1995a). Reents *et al.* (1995b) found that heritabilities for SCS defining the contemporary group as herd-test date were higher compared to those defining a contemporary group as herd-year-season of calving.

The data structure, arithmetic means and standard deviations of the selected datasets are presented in Table 3.1. SCS follows a curvilinear pattern with regards to average SCS and standard deviation as age of calving increased. SCS is also constantly higher in the summer months compared to the winter months for both breeds. Haile-Mariam *et al*. (2000) observed an increase in mean SCC and $\log_e$SCC with parity which were in agreement with that reported for Holstein-Friesian cattle in the UK by Mrode *et al.* (1998), as well as results from Pagnacco *et al.* (1994). According to Detilleux *et al.*

Table 3.1 Number of test-day observations by parity, age of calving and season of calving with their arithmetic means and standard deviations of the selected datasets for somatic cell score.

				HOLSTEINS			
	AGE OF	SEASON OF CALVING					
Parity	CALVING	April - September			October – March		
	months	number	$\overline{X}$	SD	number	$\overline{X}$	SD
1	<24	3 670	11.24	1.27	3 563	11.43	1.33
	24-26	17 770	11.31	1.26	15 823	11.40	1.26
	27-29	16 064	11.32	1.20	15 588	11.40	1.18
	30-32	10 180	11.30	1.20	9 821	11.42	1.19
	>32	10 335	11.35	1.17	8 726	11.45	1.14
2	<36	2 540	11.78	1.37	2 745	12.04	1.37
	36-38	13 086	11.72	1.36	11 617	11.91	1.37
	39-41	15 385	11.74	1.35	15 562	11.87	1.33
	42-46	17 432	11.69	1.33	18 439	11.89	1.30
	>46	10 317	11.73	1.29	8 784	11.85	1.29
3	<48	871	12.04	1.30	844	12.28	1.35
	48-56	15 655	12.09	1.41	14 552	12.18	1.35
	57-60	5 378	12.01	1.36	4 877	12.11	1.32
	>60	3 238	11.99	1.40	2 755	12.12	1.28
				JERSEYS			
1	<24	1 218	11.43	1.07	1 159	11.45	1.11
	24-26	5 718	11.31	1.09	5 045	11.43	1.10
	27-29	9 924	11.25	1.08	9 349	11.42	1.09
	30-32	11 278	11.28	1.10	10 687	11.46	1.07
	>32	16 805	11.31	1.11	16 198	11.46	1.06
2	<36	946	11.33	1.18	855	11.44	1.23
	36-38	4 024	11.15	1.19	3 440	11.43	1.24
	39-41	7 036	11.23	1.21	6 781	11.48	1.22
	42-46	13 482	11.39	1.22	13 053	11.66	1.18
	>46	9 912	11.44	1.18	9 316	11.74	1.14
3	<48	416	11.59	1.31	318	11.80	1.38
	48-56	8 471	11.52	1.30	6 863	11.75	1.32
	57-60	4 288	11.67	1.28	3 903	11.88	1.22
	>60	3 486	11.77	1.23	3 458	12.03	1.22

(1997) and Reneau (1986) the increase in log$_e$SCC with parity and age is generally attributed to the fact that older cows have a greater opportunity for exposure to mastitis-causing pathogens.

Estimated lactation curves for SCS are indicated in Figures 3.1 and 3.2, for Holstein and Jersey cows, respectively and look similar to inverted lactation curves for milk production. These curves were adjusted to a common mean and therefore do not display the increase in absolute level of SCS with increasing parity or between summer and winter seasons. These graphs, however, indicate clearly that curves for first parity differ significantly from curves of second and third parities for both breeds.

Curves from later parities increased more rapidly after the minimum in SCS was reached during the second month in milk, compared to that of the first parity. Studies of Reents *et al*. (1995b), Zhang *et al*. (1994), Schutz *et al*. (1990) and Wiggans & Shook (1987) confirm these results. Although not indicated by the graphs due to the common mean used, the Best Linear Unbiased Estimates (BLUEs) of the effect of season on SCS indicated that higher SCS occur in summer months compared to winter months (Table 3.2), but according to the graphs SCS increased more rapidly in winter than in summer months. Reents *et al.* (1995b) reported further that across age and season groups, curves differed mainly in height (higher SCS during the summer and with increasing age within parity), but were similar in shape. The shape of the curves for the second and third lactations of the Holstein cows, were almost identical after the second month in milk, while the shape of the graphs of the Jersey cows differed more between parities and seasons. Higher SCS were evident for Holstein cows in all parities and seasons compared to SCS of Jersey cows. Haile-Mariam *et al.* (2000) also showed that the effect of stage of lactation on log$_e$SCC is 'inverted' relative to milk production and that the trend for the second and third parities was nearly the same, but different from that of the first parity.

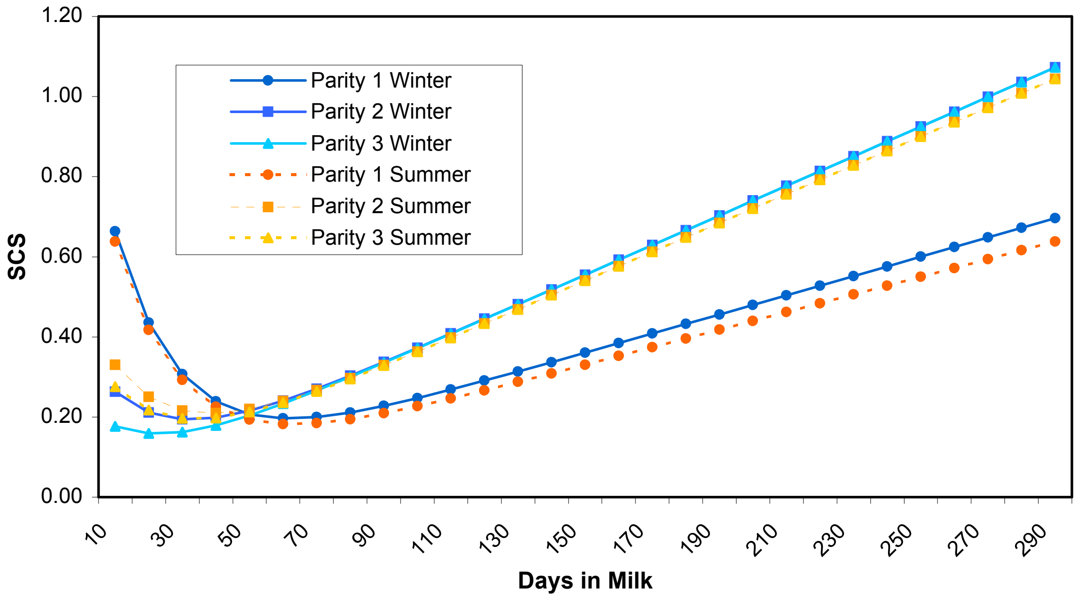


Figure 3.1 Estimated lactation curves of somatic cell score (SCS) from the Multiple Lactation Model for parity 1, 2 and 3, nested in season, for Holstein cows.

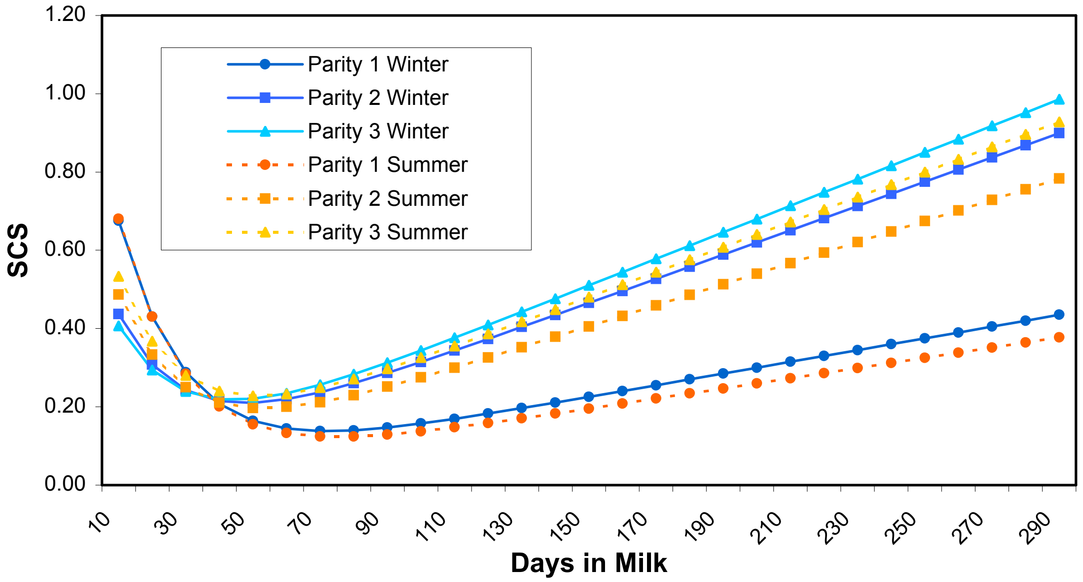


Figure 3.2 Estimated lactation curves of somatic cell score (SCS) from the Multiple Lactation Model for parity 1, 2 and 3, nested in season, for Jersey cows.

Ratios and correlations estimated for SCS with the ML and RM-models, are listed in Table 3.3 for Holstein and Jersey cows.

Table 3.2 Best Linear Unbiased Estimates (BLUEs) from the Multiple Lactation Model for the effect of season on SCS in the different lactations for Holstein and Jersey cows.

	HOLSTEINS			**JERSEYS**		
	Winter	**Summer**	**Solver**	**Winter**	**Summer**	**Solver**
Parity 1	-0.0428	0.0474	IOC	11.0768	11.1990	Jacobi
Parity 2	-0.0468	0.0504		10.9775	11.1796	
Parity 3	-0.0413	0.0462		11.2057	11.3832	

For the ML-model, a trivariate analysis was done for the Holsteins which converged after 97 iterations giving semi-positive definite matrices. For the Jerseys, however, matrices had to be constructed from bivariate analyses to end up with semi-positive definiteness. Heritabilities were low for both breeds, increasing slightly with parity number. Lactations 2 and 3 were highly correlated, with lower correlations between lactations 1 and 2 and the lowest correlations between lactations 1 and 3. The permanent environmental effect was much higher than the heritabilities for all lactations of both breeds, with low to moderate correlations between the lactations – the highest correlation again between lactations 2 and 3 of both breeds. Reents *et al.* (1995b) reported that estimates of genetic correlations for SCC between adjacent lactations vary considerably, from 0.44 to near unity and that genetic correlations tend to be lower as distance between lactations increases. They published genetic correlations of 0.88, 0.79 and 0.95 between lactations 1 and 2, 1 and 3 and 2 and 3, respectively and found that correlations between permanent environmental effects were smaller (0.29, 0.19 and 0.46 between lactations 1 and 2, 1 and 3 and 2 and 3, respectively). Haile-Mariam *et al.* (2000) reported average heritability estimates for Holstein-Friesian cows of Australia of 0.08 for lactation 1, 0.09 for

Table 3.3 Heritabilities, permanent environmental (PE) effects and correlations (± SE) of the ML- (heritabilities and PE on the diagonal, correlations above the diagonal) and RM-models for SCS of Holstein and Jersey cows.

MULTIPLE LACTATION MODEL

Heritabilities and Correlations			***PE and Correlations***		
Lactation 1	**Lactation 2**	**Lactation 3**	**Lactation 1**	**Lactation 2**	**Lactation 3**
Holsteins					
0.050±0.004	0.839±0.027	0.791±0.043	***0.279±0.004***	0.378±0.009	0.314±0.013
	0.070±0.005	0.973±0.021		***0.310±0.005***	0.531±0.011
		0.069±0.008			***0.338±0.008***
Jerseys					
0.039	0.807±0.059	0.750±0.083	***0.310***	0.413±0.011	0.236±0.018
	0.044	1.000±0.000		***0.328***	0.502±0.013
		0.046			***0.359***

REPEATABILITY MODEL

Heritability	***PE***
Holsteins	
0.191±0.003	0.190±0.002
Jerseys	
0.182±0.003	0.206±0.003

lactation 2 and 0.11 for lactation 3 from random regression sire test-day models. Reents *et al*. (1995b) stated from their literature survey that estimates of heritability for various expressions of SCC, mainly linear somatic cell score, are moderately low and Reents *et al.* (1995a) concluded from various studies that SCS is a lowly heritable trait ($h^2 = 0.10$ to 0.15). Reents *et al*. (1995b) reported estimates for Canadian Holsteins of 0.09 for lactation 1 and 2 and 0.11 for lactation 3 and estimates for PE of 0.42 for lactation 1, 0.41 for lactation 2 and 0.43 for lactation 3, using Gibbs sampling and a test-day model treating lactations as different traits. Reents *et al.* (1995a) published heritability estimates of 0.08, 0.13 and 0.14 for lactations 1, 2 and 3 respectively and estimates for permanent environmental effects of 0.38, 0.34 and 0.34 for lactations 1, 2 and 3 respectively, for German Holstein cattle, using test-day models treating lactations as different traits. Heritability estimates from the ML-model from this study are therefore on the lower range of those reported in the literature, while the permanent environmental estimates correspond better.

A redistribution of genetic and permanent environmental variances occurred when comparing estimates from the RM and ML-models. The RM-model estimated heritabilities for SCS that is 3-4 times higher compared to estimates from the ML-model : 0.19 for the Holsteins and 0.18 for the Jerseys. This happened at the expense of the permanent environmental effect, as the ML-model estimated these effects to be 4-5 times higher than the heritabilities, but the ML-model estimated these effects in the same range as the heritabilities (0.19 for the Holsteins; 0.21 for the Jerseys). The RM-model might therefore not account for all possible permanent environmental variances due to repeated measures within and across lactations.

The estimates from the RM-model are more comparable with estimates used by other countries participating in INTERBULL runs to obtain MACE breeding values. Countries using fixed regression test-day animal models for prediction of breeding values for SCS, are Estonia (h^2=12%), Italy (h^2=8% only first lactations) and Switzerland (h^2=20-31%, depending on the breed), while countries implementing random regression test-day models are Canada (h^2=25-33%), Germany (h^2=23%) and The Netherlands (h^2=35%) (http://www-interbull.slu.se).

Breeding values converged at 4 313 iterations for the Holsteins (± 12 hours) and 3 296 iterations (± 3 hours) for the Jerseys, using the RM-model. The ML-model is of course more computationally demanding. After 20000 rounds of iteration, animal effects were converged (first derivative is zero) for lactations 1 and 2 at a stopping criteria of 0.0001 (standardized maximum change), while those of lactation 3 were still at 0.00392 for the Holsteins (±168 hours) and 0.00311 for the Jerseys (± 53 hours). Rank correlations between the breeding values of the RM-model and the breeding value indices of the ML-model of sires with at least 50 daughters in 10 herds (proven sires), as well as for less reliable sires (less than 50 daughters in 10 herds) are indicated in Table 3.4.

Table 3.4 Spearman correlation coefficients between breeding values from the RM-model and breeding value indices of the ML-model (MLeq for equal weights and MLweig for 0.25*lactation 1 + 0.65*lactation2 + 0.10*lactation 3) of sires with at least 50 daughters in 10 herds (proven sires), as well as for sires with less than 50 daughters in 10 herds (less reliable sires) for Holstein and Jersey cattle.

	HOLSTEINS			**JERSEYS**		
	Proven Sires					
	No	**MLeq**	**MLweig**	**No**	**MLeq**	**MLweig**
RM	334	0.977	0.976	118	0.972	0.971
	Less Reliable Sires					
	2 169	0.892	0.894	1 686	0.861	0.858

Construction of a breeding value index, using equal weights (MLeq) for the three lactations for breeding values from the ML-model, gave almost the same ranking correlations for proven as well as less reliable sires with breeding values from the RM-models, as indices weighing lactations by 0.25, 0.65 and 0.10 for lactations 1, 2 and 3, respectively (MLweig). Proven sires rank almost the same, regardless of the model used, while 14% of the Jersey and 11% of the Holstein sires with less than 50 daughters in 10 herds, will change places in the rank between the different models.

3.5 CONCLUSIONS

Correlations between lactations are not in unity, indicating that lactations should be treated as different traits, nevertheless the RM-model estimates more competitive variances and requires extensively less computing time in terms of national and international genetic evaluations. Although ML-models estimate more accurate breeding values, enabling breeders to weigh information from different lactations, only minor changes in the ranking of proven sires occurred in the comparison of breeding values from the ML- and RM-models.

3.6 REFERENCES

Boettcher, P.J., Hansen, L.B. Van Raden, P.M. & Ernst, C.A., 1992. Genetic evaluation of Holstein bulls for somatic cells in milk of daughters. J. Dairy Sci. 75, 1127-1137.

Boichard, D. & Rupp, R., 2001. Phenotypic and genetic relationships between somatic cell counts and clinical mastitis in French Holstein cows. Interbull Bulletin 26, 66-72.

Detilleux, J., Leroy, P. & Volxckaert, D., 1997. Alternative use of somatic cell counts in genetic selection for mastitis resistance. Interbull Bull. Uppsala, Sweden 15, 34-44.

Emanuelson, U., Danell, B. & Rupp, R., 1988. Genetic parameters for clinical mastitis, somatic cell counts and milk production estimated by multiple-trait restricted maximum likelihood. J.Dairy Sci. 71, 467-476.

Emanuelson, U. & Perrson, E., 1984. Studies on somatic cell counts in milk from Swedish dairy cows. I. Non-genetic causes of variation in monthly test-day results. Acta Agric. Scand. 34, 33-44.

Groeneveld, E. & Garcia-Cortes, A., 1998. VCe4.0, a (co)variance component package for frequentists and Bayesians. Proc. 6th WCGALP. 27, 455-456.

Groeneveld, E. & Kovac, M., 1990. A generalised computing procedure for setting up and solving mixed linear models. J. Dairy Sci. 73, 513-531.

Haile-Mariam, M., Goddard, M.E. & Bowman, P.J., 2000. Estimates of Genetic Parameters for Daily Somatic Cell Count of Australian Dairy Cattle. J.Dairy Sci. 84, 1255-1264.

Harmon, R.J., 1994. Physiology of mastitis and factors affecting somatic cell counts. J. Dairy Sci. 77, 2103-2112.

Kadarmideen, H.N. & Pryce, J.E., 2001. Genetic and economic relationships between somatic cell count and clinical mastitis and their use in selection for mastitis resistance in dairy cattle. Anim. Sci. 73, 19-28.

Kennedy, B.W., Sethar, M.S., Tong, A.K.W., Moxley, J.E. & Downey, B.R., 1982. Environmental factors influencing test-day somatic cell count in Holsteins. J. Dairy Sci. 65, 275-280.

Mostert, B.E., Theron, H.E. & Kanfer, F.H.J., 2001. The effect of calving season and age at calving on production traits of South African dairy cattle. S. Afr. J. Anim. Sci. 31, 205-214.

Mrode, R.A. & Swanson, G.J.T., 1996. Genetic and statistical properties of somatic cell count and its suitability as an indirect means of reducing the incidence of mastitis in dairy cattle. Anim. Breed. Abstr. 64, 847-857.

Mrode, R.A. , Swanson, G.J.T. & Winters, M.S., 1998. Genetic parameters and evaluation for somatic cell counts and its relationship with production and type traits in some dairy breeds in the United Kingdom. Anim. Sci. 66, 569-576.

Nash, D.L., Rogers, G.W., Cooper, J.B., Hargrove, G.L., Keow, J.F. & Hansen, L.B., 2000. Heritability of clinical mastitis incidence and relationships with sire transmitting abilities for somatic cell score, udder type traits, productive life and protein yield. J.Dairy Sci. 83, 2350-2360.

Pagnacco, G. Miglior, F., Zhang, W.C., Dekkers, J.C.M. & Burnside, E.B., 1994. Genetic evaluation for somatic cell count and relationship with inbreeding in Canadian Holsteins. Proc. 5th World Congr. Genet. Appl. Livest. Prod., Guelph, Ontario, Canada 17, 93-96.

Reents, R., Dekkers, J.C.M. & Schaeffer, L.R., 1995a. Genetic evaluation for somatic cell score with a test day model for multiple lactations. J. Dairy Sci. 78, 2858-2870.

Reents, R. & Dopp, L., 1996. Genetic evaluation for dairy production traits with a test day model for multiple lactations. Proc. Open Session of the Interbull Annual Meeting, Veldhoven, The Netherlands, June 23-24 1996, 113-117.

Reents, R., Jamrozik, J., Schaeffer, L.R. & Dekkers, J.C.M, 1995b. Estimation of genetic parameters for test day records of Somatic Cell Score. J.Dairy Sci., 78, 2847-2857l.

Reneau, J.K., 1986. Effective use of dairy herd improvement somatic cell counts in mastitis control. J. Dairy Sci. 69, 1708-1720.

Rogers, G.W., Banos, G., Sander Nielsen, U. & Philipsson, J., 1998. Genetic correlations among somatic cell scores, productive life and type traits from the United States and udder health measures from Denmark and Sweden. J. Dairy Sci. 81, 1445-1453.

SAS, 1996. SAS user's guide : Statistics, Release 6.12. SAS Institute Inc., Cary, North Carolina, USA.

Schutz, M.M., Hansen, L.B., Steuernagel, G.R. & Kuck, A.L., 1990. Variation of milk, fat, protein, and somatic cells for dairy cattle. J. Dairy Sci. 73, 484-493.

Sethar, M.S., Kennedy, B.W., Tong, A.K.W., Moxley, J.E. & Downey, B.R., 1979. Genetic and environmental factors affecting test day somatic cell counts in Holsteins. J. Dairy Sci. 62 (Suppl. 1), 48 (Abstr).

Simianer, H., Solbu, H. & Schaeffer, L.R., 1991. Estimated genetic correlation between disease and yield traits in dairy cattle. J. Dairy Sci. 74, 4358-4365.

Wiggans, G.R. & Shook, G.E., 1987. A lactation measure of somatic cell count. J. Dairy Sci. 70, 2666-2672.

Wilmink, J.B.M., 1987. Adjustment of test-day milk, fat and protein yield for age season and stage of lactation. Livest. Prod. Sci. 16, 335-348.

Zhang, W.C., Dekkers, J.C.M., Banos, G. & Burnside, E.B., 1994. Adjustment factors and genetic evaluation for somatic cell score and relationships with other traits of Canadian Holsteins. J. Dairy Sci. 77, 659-665.

CHAPTER 4

TEST-DAY MODELS FOR PRODUCTION TRAITS IN DAIRY CATTLE IN SOUTH AFRICA

B.E. Mostert#, H.E. Theron, E. Groeneveld[1] & F.H.J. Kanfer[2]
ARC-AII, Private Bag X2, Irene, 0062, South Africa.
[1] FAL, Institut fur Tierzucht, 31535, Neustadt, Germany
[2]Department of Statistics, University of Pretoria, Pretoria, 0002, South Africa.

4.1 ABSTRACT

Methodology has been developed to implement test-day models in the national genetic evaluation of dairy breeds for production traits in South Africa. Semi-positive definite covariance matrices have been estimated, using multitrait, multi-lactation, fixed regression test-day BLUP animal models, including test-day records of the first three lactations as repeated measures for the Holstein and Jersey breeds. Heritability estimates were 0.40 for milk yield, 0.25 for butterfat yield and 0.37 for protein yield for the Holstein breed and 0.39 for milk yield, 0.21 for butterfat yield and 0.34 for protein yield for the Jersey breed. These estimates are well in the range of estimates reported by countries participating in international evaluations done by INTERBULL (International Bull Evaluation Service) using test-day models.

Keywords : Fixed regressions, Holstein, Jersey, Wilmink curve
#Corresponding author. E-mail : bernice@arc.agric.za

4.2 INTRODUCTION

Dairy sires in South Africa were first genetically evaluated through progeny groups, using data recorded by the National Livestock Improvement Scheme. This led to the utilization of contemporary comparison methods to estimate the breeding values of sires. Dairy animals received breeding values from BLUP (Best Linear Unbiased Prediction) Methodology for the first time in 1987, when breeding values were estimated, using a Sire Model. In 1992 the Animal Model was fitted to dairy records to estimate breeding values in single trait analyses. Since 1999 multitrait analyses were developed for the South African dairy breeds, where 305-day milk, butterfat and protein first lactation yields were evaluated together in order to utilize the genetic correlations between the traits to estimate breeding values even more accurately (Loubser *et al*., 2001). Genetic groups were also incorporated into the pedigrees in order to qualify in 2000 for participation in international evaluations done by INTERBULL (International Bull Evaluation Service in Sweden) for the estimation of MACE (Multiple Across Country Evaluation) breeding values.

Using only completed first lactation records in genetic analyses has of course several disadvantages. A cow that did not participate in milk recording during her first lactation or for some reason her first lactation was terminated before 240 days in milk, never received a breeding value based on her own performance, regardless of her performance in later lactations. Furthermore, the performance of the cows in second and third lactations is never included, rendering less accurate genetic evaluations especially for “late bloomers”.

The basis of a 305-day yield is a set of test-day yields taken approximately every 35 days in milk in South Africa. Incomplete lactations are extended to a 305-day basis following a set of well-defined rules (Ptak & Schaeffer, 1993). Such projection procedures usually assume a fixed shape of lactation curves for cows calving in the same season, of the same age class and herd level and tend to underestimate 305-day yields from early test-days for more persistent cows and overestimate yields for the less persistent cows (Mrode *et al*., 2002). One way to avoid the problem of extension of test-day yields into a 305-day record, would be to use test-day yields for genetic evaluation of dairy sires and cows, rather than 305-day yields. The aim of this study was therefore to develop a genetic model and to estimate covariance matrices that could be used to evaluate sires and cows in South Africa, based on test-day yields from lactations 1 to 3.

4.3 MATERIALS AND METHODS

Data consisted of test-day records for milk, butterfat and protein yield for lactations 1, 2 and 3 of Holstein and Jersey cows calving from 1995 to 2002. These records were downloaded from the INTERGIS (Integrated Registration and Genetic Information System) for cows participating in the South African Dairy Animal Improvement Scheme. Basic edits included deletion of records with unknown herds, unknown birthdates, test-days recorded before six days in milk or after 305 days in milk, records of crossbred cows and age restrictions within lactations to ensure reasonable calving ages in a specific lactation In order to ensure a well-linked data structure for variance component estimation, the following data selections were carried out :

Only records with both parents known, were used. Contemporary groups were then defined as Herd x Test-date x Parity x Number of milkings/day. 33% of these contemporary groups of the first parity were randomly selected for the Holstein and 60% for the Jersey analyses. The remaining records were subjected to the following requirements :

- Cows included in the analysis should have a first parity
- There should be daughters of at least two sires in a contemporary group
- The contemporary group should consist of at least five records
- Each sire should be represented in at least three contemporary groups
- Each lactation was divided into 10 stages of 30 days each, except for the 10th stage which consisted of 35 days (270-305). Only one test-day record per stage per cow was included in the analyses.
- A cow should have at least three test-days per lactation to be selected.
- A sire should have at least six daughters in the first lactation.

This selection was done based on the study of Haile-Mariam *et al.* (2000).

After research was done on the estimation of breeding values for somatic cell count using test-day models in South Africa (Mostert *et al*., 2004), it was decided to include records from the first three lactations as repeated measures of the first lactation, assuming a genetic correlation of 1 across lactations. By using this methodology, it was still possible in terms of computer capacity and time, to do multitrait analyses for production traits. This was a prerequisite for developing a new genetic model based on test-days, as the South African Dairy Animal Improvement Scheme has allowed breeders to only record milk yields since 2001 (Dr Japie van der Westhuizen, 2004: Personal communication, INTERGIS Manager, e-mail: japie@arc.agric.za). Thus, the use of genetic correlations to estimate breeding values for protein and butterfat yields became vitally important in genetic evaluations of dairy breeds in South Africa. Starting with the Holstein data a subset of 34 662 test-day records were used from the abovementioned selection for variance component estimation. This included 10 955 test-day records for lactation 1, 14 955 records for lactation 2 and 8 752 for lactation 3. This is the data of 3 093 cows, representing 250 sires and 3 052 dams. For the Jerseys the selection ended with 88 572 test-day records for the first lactation, 70 009 records in the second lactation and 30 870 records in the third lactation, representing 17 758 cows, 473 sires and 14 941 dams.

(Co)variance components were estimated using VCE4 (Groeneveld & Garcia-Cortes, 1998) with the following genetic model :

$$y_{ijklmnp} = \mu + HTDLM_{im} + A_j + PE_{jm} + S_{km} + AC_{lm} + wilmink(S_{km}) + CI_{jm} + e_{ijklmnp}$$

Where

$y_{ijklmnp}$ = p^{th} test-day milk, butterfat or protein yield of cow j in lactation m,
μ = mean yield
$HTDLM_{im}$ = fixed effect of herd x test-date x parity x number of milkings group
A_j = animal additive genetic effect
PE_{jm} = permanent environmental (random) effect within lactation to account for common effects of environment associated with all test-day records of cow j in lactation m
S_{km} = fixed effect of calving season in lactation m
$\mathrm{AC}_{\mathrm{lm}}$ = fixed effect of ageclass in lactation m
$wilmink(S_{km})$ = Wilmink curve (Wilmink, 1987) modelled on days in milk within season in lactation m (regression)
$\mathrm{CI}_{\mathrm{n}m}$ = fixed effect of calving interval class in lactation m
$e_{ijklmnp}$ = random residual error

Season was defined as winter (April – September) *versus* summer (October – March), while the same ageclasses were allocated as in the derivation of standard lactation curves by Mostert *et al.* (2001). Calving interval classes were allocated using standard deviation units.

A series of univariate and bivariate analyses were run as well as a trivariate analysis for each breed.

4.4 RESULTS AND DISCUSSION

The advantages of a test-day model include a more accurate correction for environmental effects relevant to each test-day record (Ptak & Schaeffer, 1993); the ability to model the shape of the lactation curve to differ for each cow; estimation of persistency evaluations (Jamrozik *et al.*, 1997); as well as the use of early predictors of genetic merit for selection decisions, thereby decreasing the generation interval. Other advantages are that records do not need to be extended, erasing debates over extension factors; better modeling of the effect of pregnancy can occur because a test-day can directly be linked to whether or not a cow is pregnant and that further subdivision of the contemporary group into management groups is feasible, if they are recorded (Swalve, 2000). Test-day models are of course computationally very demanding. For evaluations on a national scale, the size of the equation system can go to hundreds of millions of equations, depending on the size of the database and the specific model defined. Furthermore, all the individual test-day yields on every cow need to be stored.

Table 4.1 Arithmetic means and standard deviations of the selected datasets.

	HOLSTEIN			**JERSEY**		
	No records	**Mean**	**SD**	**No Records**	**Mean**	**SD**
Milk	34 662	27.01	9.36	189 449	16.63	5.22
Butterfat	34 553	0.95	0.37	189 153	0.76	0.26
Protein	34 553	0.85	0.28	189 155	0.60	0.18

Table 4.1 indicates the arithmetic means and standard deviations of the selected datasets. Holstein cattle produced on average 10.4 kg milk, 0.19 kg butterfat and 0.25 kg protein more per day compared to Jersey cattle and also showed more variation for all traits.

All analyses converged with status 1 (defined as standardized maximum change of the solutions from one round to the next with a stopping criterion of 0.0001), giving semi-positive definite matrices. Table 4.2 indicates heritabilities and genetic correlations obtained from the trivariate analyses. Heritabilities were slightly higher for the Holstein breed for all traits, while the genetic correlation between milk and butterfat yields was similar (84%) and between milk and protein yields was slightly higher for the Holstein breed (94% *vs* 92%) compared to the Jersey breed.

Table 4.2 Heritabilities (on the diagonal) and correlations (above the diagonal) for milk, butterfat and protein yield (kg/day) of Jersey and Holstein cattle.

	HOLSTEIN			**JERSEY**		
	Milk	**Butterfat**	**Protein**	**Milk**	**Butterfat**	**Protein**
			DIRECT			
Milk	**0.40 ± 0.007**	0.84 ± 0.006	0.94 ± 0.002	**0.39 ± 0.004**	0.84 ± 0.002	0.92 ± 0.001
Butterfat		**0.25 ± 0.006**	0.88 ± 0.005		**0.21 ± 0.002**	0.91 ± 0.002
Protein			**0.37 ± 0.006**			**0.34 ± 0.002**
			PERMANENT ENVIRONMENT			
Milk	**0.13 ± 0.008**	0.98 ± 0.005	0.99 ± 0.002	**0.23 ± 0.003**	0.96 ± 0.002	0.98 ± 0.001
Butterfat		**0.09 ± 0.006**	0.99 ± 0.004		**0.16 ± 0.002**	0.96 ± 0.002
Protein			**0.11 ± 0.007**			**0.20 ± 0.002**
			RESIDUAL			
Milk	**0.47 ± 0.006**	0.67 ± 0.002	0.93 ± 0.001	**0.39 ± 0.002**	0.64 ± 0.001	0.91 ± 0.00
Butterfat		**0.67 ± 0.007**	0.66 ±0.003		**0.63 ± 0.002**	0.62 ± 0.001
Protein			**0.53 ± 0.007**			**0.46 ± 0.002**

Only the genetic correlation between butterfat and protein yields was higher for the Jersey breed compared to the Holstein breed (91% *vs.* 88%). The correlations between the permanent environmental effects were in the high nineties for both breeds, while the direct estimates for permanent environment for the Holstein breed were almost half of that of the Jerseys for all traits. Residual estimates (direct and correlations) were higher for the Holsteins compared to the Jerseys for all traits. Both breeds showed the same pattern regarding residual correlations, namely correlations in the sixties between milk and butterfat yields and between butterfat and protein yields, whereas the correlations between milk and protein yields were in the nineties for both breeds.

These estimates are within the range of estimates reported by countries participating in INTERBULL using test-day methodology (www-interbull.slu.se) (Table 4.3). However, heritabilities can be defined in many different ways, especially for test-day models and a comparison only makes sense if heritabilities are defined the same way. For countries like Germany, The Netherlands and Canada, the heritability estimates provided for INTERBULL evaluations is expressed on a daily basis and combined lactation basis. These heritability estimates look therefore rather high, although their original estimates on a daily basis, are lower. The heritability from a fixed regression test-day model reflects the heritability of a single test-day observation, whereas the heritability from a lactation model reflects the heritability of an average of several test-day observations.

Table 4.3 Specifications of models used and heritabilities reported by countries participating in INTERBULL evaluations using test-day methodology.

Country	Model	Milk	Butterfat	Protein
Canada	MT-ML-RR-TD	0.36-0.39	0.33-0.37	0.35-0.37
Estonia	ST-ML-FR-TD	0.27	0.23	0.24
Germany	ST-ML-RR-TD	0.49	0.48	0.48
NLD/BEL	ST-ML-RR-TD	0.59	0.58	0.52
Switzerland	ST-ML-FR-TD	0.36	0.30	0.32

MT=Multitrait ML=Multi-lactation RR=Random Regression TD=Test-day ST=Single trait FR=Fixed Regression

It is therefore expected that the heritability of a lactation model will be higher than that of a fixed regression test-day model (Dr Freddy Fikse, 2006: Personal communication, INTERBULL Centre Director, e-mail: Freddy.Fikse@hgen.slu.se). Heritability estimates obtained with this methodology were however, higher for milk and protein yield compared to estimates from the previously used Lactation Model. Heritability estimates on a daily basis obtained with this methodology were therefore high and the portion of variance due to permanent environmental effects was low, while the genetic and permanent environmental correlations were also rather high among the three production traits.

4.5 CONCLUSIONS

Methodology has been developed to implement test-day models in the national genetic evaluation of dairy breeds for production traits in South Africa. Semi-positive definite covariance matrices have been estimated, using multitrait, multi-lactation, fixed regression test-day models, including test-day records of the first three lactations as repeated measures for the Holstein and Jersey breeds. These methods will be implemented after participation in the test-runs of INTERBULL in Sweden during March and September 2004.

4.6 REFERENCES

Groeneveld, E. & Garcia-Cortes, A., 1998. VCe4.0, a (co)variance component package for frequentists and Bayesians.Proc. 6th WCGALP. 27, 455-456.

Haile-Mariam, M., Goddard, M.E. & Bowman, P.J., 2000. Estimates of Genetic Paramters for Daily Somatic Cell Count of Australian Dairy Cattle. J. Dairy Sci. 84, 1255-1264.

Jamrozik, J., Dekkers, J.C.M. & Schaeffer, L.R., 1997. Genetic evaluation of dairy cattle using test day yields and random regression models. J. Dairy Sci. 80, 1217-1226.

Loubser, L.F.B, Banga, C.B., Scholtz, M.M. & Hallowell, G.J. (Eds.), 2001. Dairy Herd Improvement in South Africa. The use of breeding values in the South African Dairy Industry. 25.

Mostert, B.E., Theron, H.E. & Kanfer, F.H.J., 2001. The effect of calving season and age at calving on production traits of South African dairy cattle. S. Afr. J. Anim. Sci. 31, 205-214.

Mostert, B.E., Banga, C.B., Theron, H.E., Groeneveld, E. & Kanfer, F.H.J., 2004. Breeding value estimation for somatic cell score in South African dairy cattle. Proc. 2nd Joint Congr. Grassland Soc. Of South Africa and SASAS. 28 June-1 July 2004, Worcestor.

Mrode, R.A., Swanson, G.J.T. & Lindberg, C.M., 2002. Efficiency of part lactation test day records for genetic evaluations using fixed and random regression models. Anim. Sci. 74, 189 – 197.

Ptak, E. & Schaeffer, L.R., 1993. Use of test day yields for genetic evaluation of dairy sires and cows. Livest. Prod. Sci. 34, 23-34.

Swalve, H.H., 2000. Symposium: Test-day Models. Theoretical basis and computational methods for different test-day genetic evaluation methods. J. Dairy Sci. 83, 1115-1124.

Wilmink, J.B.M., 1987. Adjustment of test-day milk, fat and protein yield for age season and stage of lactation. Livest. Prod. Sci. 16, 335-348.

CHAPTER 5

TEST-DAY MODELS FOR SOUTH AFRICAN DAIRY CATTLE FOR PARTICIPATION IN INTERNATIONAL EVALUATIONS

B.E. Mostert#, H.E. Theron, F.H.J. Kanfer[1] & E. van Marle-Köster[2]
ARC-LBD, Private Bag X2, Irene, 0062, South Africa
[1]Department of Statistics, University of Pretoria, Pretoria, 0002, South Africa
[2]Department of Animal and Wildlife Sciences, University of Pretoria, Pretoria, 0002, South Africa

---

5.1 ABSTRACT

Variance components and breeding values of production traits and somatic cell score of South African Guernsey, Ayrshire, Holstein and Jersey breeds have been estimated using a multi-lactation repeatability test-day model, including tests of the first three lactations as repeated measures and fitting the permanent environmental effect across lactations. Multitrait evaluations were done for the production traits (milk, butterfat and protein) and single trait evaluations for somatic cell score. Heritability estimates were comparable with yield and somatic cell score estimates obtained by test-day models from other countries (17-24% for milk yield; 10-13% for butterfat yield; 14-19% for protein yield and 6-8% for somatic cell score). Proofs of qualifying sires were sent to INTERBULL (International Bull Evaluation Service) for participation in the March 2005 test-runs. Genetic correlations between South Africa and other participating countries, estimated by INTERBULL, compared well with those amongst the other participating countries. Trend validation tests were successful using this methodology for all traits and breeds except for somatic cell score of the Guernsey breed, due to insufficient data for this trait. South Africa can now participate in routine INTERBULL evaluations to obtain MACE (multiple across country evaluation) breeding values, using this methodology.

Keywords : Ayrshire, Guernsey, Holstein, Jersey, production traits, repeatability model, somatic cell score

#Corresponding author. E-mail : bernice@arc.agric.za

5.2 INTRODUCTION

Estimates of genetic and phenotypic parameters and the various components of variances and covariances therein are necessary elements for the estimation of breeding values. These estimates need to be relevant for each population and the actual situation where they are to be used (Danell, 1982). In recent years, interest in test-day records has increased among dairy geneticists and among members of the dairy cattle breeding industry (Swalve, 2000), *i.e.* the inclusion of test-day records as they are in genetic evaluations instead of using lactation records that are calculated from them (Swalve, 1998). A test-day model for genetic evaluation can account for factors that are specific to each test-day, such as management groups within a herd on a test-day, day of the year (including weather conditions), and, for each cow, days in milk, pregnancy status, medical treatments and number of times milked on the test-day. Many of these factors change for a cow from one test-day to the next and would be difficult to model for 305-day yields (Jamrozik *et al*., 1997).

Test-day models have been implemented for South African Holstein and Jersey cows, including test-days of the first three lactations as repeated measures and fitting the permanent environmental effect within lactation (Mostert *et al*., 2004a,b). The Jersey evaluation was submitted to the March 2004

INTERBULL test-run. INTERBULL (International Bull Evaluation Service) is the sub-committee of ICAR (The International Committee for Animal Recording) who is responsible for global genetic evaluations. Genetic correlations amongst countries participating in INTERBULL evaluations are estimated by INTERBULL during these test-runs, for subsequent estimation of MACE (multiple across country evaluation) breeding values in international routine evaluations. Data used by INTERBULL for estimation of genetic correlations are de-regressed breeding values for all AI bulls that have daughters in at least 10 herds in the country of origin. Correlations are estimated using the software package developed at Holstein Association USA, by Klei & Weigel (1998). Correlations are estimated simultaneously for all countries, except for Holstein where subsets of countries are considered.

Although South Africa's test-day methodology was accepted after the March 2004 test-run, concern was expressed regarding low correlations of South Africa with countries sharing a reasonable number of common sires, especially for protein yield (eg. As low as 58% for milk, 30% for butterfat and 19% for protein yield) (Dr Freddy Fikse, 2006: Personal communication, INTERBULL Centre Director, e-mail: Freddy.Fikse@hgen.slu.se). As the permanent environmental effect accounts for environmental similarities between test-day records of the same cow, ranking of sires might be influenced when permanent environmental variance is fitted across lactations opposed to only within lactation, because different sires are linked to different numbers of test-day records per daughter per lactation. If test-day records of cows are therefore not linked across lactations, ranking of sires might be influenced as the effect of ignoring environmental similarities across lactations, will differ amongst sires. Also, fitting the permanent environmental effect across lactations, is done by most countries where records of different lactations are included as repeated measures. Harmonisation of South Africa's genetic methodology with those of other countries might therefore improve genetic correlations between South Africa and the other countries participating in INTERBULL evaluations. The aim of this study was therefore to investigate the effect of fitting the permanent environmental effect across lactations in the South African Test-Day Model for genetic evaluation of production traits and somatic cell score for the Ayrshire, Guernsey, Holstein and Jersey populations, with regards to variance components estimated and genetic correlations obtained with countries genetically linked to South Africa for successful participation in international genetic evaluations.

5.3 MATERIALS AND METHODS

For the Ayrshire, Guernsey, Holstein and Jersey breeds, data consisted of test-day records for somatic cell count and milk, butterfat and protein yield for lactations 1 - 3 calving from 1982 (1984 for the Guernseys) to 2004. These records were obtained from the INTERGIS (Integrated Registration and Genetic Information System) for cows participating in the South African Dairy Animal Improvement Scheme. The number of test-day records available was 418 806 for the Ayrshires, 179 466 for the Guernseys, 2 961 647 for the Jerseys and 8 826 442 for the Holsteins. Basic edits included deletion of records with unknown herds, unknown birth dates and calving dates, test-days recorded before five days in milk or after 305 days in milk, records of crossbred cows and age restrictions within lactations to ensure reasonable calving ages in a specific lactation (20 – 42 months for lactation 1, 30 – 54 months for lactation 2 and 40 – 66 months for lactation 3 for Holsteins and 17 – 40 months for lactation 1, 29 – 53 months for lactation 2 and 41 – 67 months for lactation 3 for the other breeds). Protein yield were set to zero for records where protein percentage was greater than 6% or less than 2%. The same was done with butterfat yield where butterfat percentage was higher than 9% or lower than 2%. Records where test-day milk yield was higher than 70kg (90 kg for Holstein) or lower than 1kg (2.5kg for Holstein) were also deleted. Lactations should fit the following requirements to be included (specifications from IRIS - the national dairy management system) :

1. First test of a lactation should be within the first 63 days in milk.
2. No interval longer than 100 days between tests of a lactation are allowed.
3. Only one interval between 60 and 100 days allowed per parity.

Lactations ending before 60 days in milk were also discarded.

Records deleted amounted to 21% of the Ayrshire, 11% of the Guernsey, 13% of the Jersey and 16% of the Holstein databases. The remaining records were used to estimate BLUP breeding values for production traits and somatic cell score.

In order to ensure a well-linked data structure for variance component estimation, the data sets were selected based on the study of Haile-Mariam *et al.* (2000). Only records with both parents known were used. Contemporary groups were defined as Herd x Test-Date x Parity x Number of Milkings/day. 25% of these contemporary groups of the first parity were randomly selected for the Holstein and 60% for the Jersey analyses. The records were further subjected to the following requirements :

- Cows included in the analysis should have a first parity.
- There should be daughters of at least three sires in a contemporary group (two for the Guernseys and Holsteins).
- The contemporary group should consist of at least five records (six for the Ayrshires).
- Each sire should be represented in at least four contemporary groups (three for Jerseys, two for Holsteins).
- Each lactation was divided into 10 stages of 30 days each, except for the 10th stage which consisted of 35 days (270-305). Only one test-day record per stage per cow was included in the analyses.
- A cow should have at least eight test-days per lactation to be selected (six for the Guernseys and Jerseys and five for the Holsteins).
- A sire should have at least six daughters in the first lactation (three for the Guernseys).

This selection was an interactive process, with the number of records available and data quality for each dataset leading to certain decisions with regards to the abovementioned requirements.

Somatic Cell Count (1000's/ml) (SCC) was transformed to log(SCC) for each test-day to achieve normality. This will be referred to as Somatic Cell Score (SCS).

(Co)variance components were estimated using VCE4 (Groeneveld & Garcia-Cortes, 1998) with the model as described in Mostert *et al*. (2004a,b) :

$$y_{ijklmnp} = \mu + HTDLM_{im} + A_j + PE_j + S_{km} + AC_{lm} + wilmink(S_{km}) + CI_{jm} + e_{ijklmnp}$$

Where

$y_{ijklmnp}$ = pth test-day milk, butterfat or protein yield of cow j in lactation m in herd x test-date x parity x milkings group i, of season k, age class l and calving interval class n

μ = mean yield

$HTDLM_i$ = fixed effect of herd x test-date x parity x number of milkings group

A_j = animal additive genetic effect

PE_j = permanent environmental (random) effect to account for common effects of environment associated with all test-day records of cow j

S_{km} = fixed effect of calving season in lactation m

AC_{lm} = fixed effect of ageclass in lactation m

$wilmink(S_{km})$ = Wilmink curve (Wilmink, 1987) modeled on days in milk within season k and in lactation m (regression)

CI_{nm} = fixed effect of calving interval class in lactation m

$e_{ijklmnp}$ = random residual error

Two seasons were defined : April – September and October – March, while the same age classes were allocated as in the derivation of standard lactation curves by Mostert *et al.* (2001). Calving interval classes were allocated using standard deviation units, as follows :

Code 1 : All records of Lactation 1.

Code 2 : All records of Lactation 2 with unknown calving intervals (no Lactation 1 records available).
Code 3 : All records of Lactation 2 where the calving interval between Lactations 1 and 2 was less than -1 standard deviation unit from the average calving interval.
Code 4 : All records of Lactation 2 where the calving interval between Lactations 1 and 2 was within 1 standard deviation unit from the average calving interval.
Code 5 : All records of Lactation 2 where the calving interval between Lactations 1 and 2 was greater than 1 standard deviation unit from the average calving interval.
Code 6 : All records of Lactation 3 with unknown calving intervals (no Lactation 2 records available).
Code 7 : All records of Lactation 3 where the calving interval between Lactations 2 and 3 was less than -1 standard deviation unit from the average calving interval.
Code 8 : All records of Lactation 3 where the calving interval between Lactations 2 and 3 was within 1 standard deviation unit from the average calving interval.
Code 9 : All records of Lactation 3 where the calving interval between Lactations 2 and 3 was greater than 1 standard deviation unit from the average calving interval.

As calving interval class did not have a significant influence ($P<0.05$) on somatic cell count, it was not included in the model for somatic cell score for variance component estimation.

This model therefore assumed consecutive test-day samples to be repeated observations on the same trait and permanent environmental effects accounted for environmental similarities between different test-days, within and across lactations.

For variance component estimation pedigrees were traced back for three generations. Multitrait analyses were run for the production traits and single trait analyses for SCS for each breed.

PEST (Groeneveld & Kovac, 1990) was used to estimate breeding values, using the unselected data set and the same models as for variance component estimation. The pedigrees were, however, traced back as far as possible and genetic groups were incorporated to ensure that base animals enter the evaluation on the appropriate genetic level. Proofs of qualifying sires were sent to INTERBULL for participation in the March 2005 test-runs.

5.4 RESULTS AND DISCUSSION

The data structure and statistics of the selected datasets are presented in Table 5.1 and 5.2. The average sizes of the contemporary groups involved in variance component estimation for the production traits, are worth mentioning. The Guernsey data set yielded the smallest contemporary group sizes of 9.3 cows per group, in comparison with the Jersey (12.4 cows) and Holstein (13.5 cows) data sets. The Ayrshire data set had an average contemporary group size of 25.8 cows and also yielded the most test-day records per cow per lactation. In the Ayrshire population, only 13 herds (29% of the herds

Table 5.1 Data structure of the selected datasets for Ayrshire, Guernsey, Holstein and Jersey breeds. First line contains number of cows, sires etc. and second line indicates average number of daughters per sire etc. in brackets.

TRAIT	COWS	SIRES (avg)	DAMS (avg)	CG (avg)	HERDS	PEDIGREE
			AYRSHIRES			
Production	7 042	281	5 264	3 814	64	13 098
		(25.1)	(1.34)	(25.9)		
SCS	3 552	207	2 736	1 245		13 098
		(17.2)	(1.30)	(32.7)		
			GUERNSEYS			
Production	3 042	266	2 466	4 412	49	6 813
		(11.4)	(1.23)	(9.3)		
SCS	2 327	192	1 729	3 310		6 813
		(12.1)	(1.35)	(8.2)		

			HOLSTEINS			
Production	14 856	1 565 (9.5)	11 975 (1.24)	6 583 (13.5)	86	44 070
SCS	4 135	737 (5.7)	3 644 (1.13)	1 416 (15.6)		44 070
			JERSEYS			
Production	5 264	505 (10.4)	4 032 (1.31)	5 747 (11.3)	90	12 713
SCS	2 007	254 (7.9)	1 630 (1.23)	1 161 (14.33)		12 713

CG = contemporary group levels SCS = Somatic Cell Score

participating in milk recording) consisted of more than 100 animals. Discipline of milk recording and therefore data quality, led to selection of data from the larger herds for the Ayrshire population. However, these herds did account for 71% of the data and are therefore still representative of the breed.

As expected, average production was highest for Holstein cows with regards to milk and protein yield per day (Table 5.2). The Ayrshires, however, produced on average more butterfat per day compared to the Holsteins, but the maximum butterfat yield per day was delivered by a Holstein cow (3.22kg/day). The Holstein cows also had the highest average SCS of 11.75 and showed the most variation when compared to the other breeds, while the average SCS of the Ayrshires was lowest of all the breeds.

Table 5.2 Statistics of the selected datasets for Ayrshire, Guernsey, Holstein and Jersey breeds.

Trait	**TD / lac**	**Avg / cow**	**Min**	**Max**	**Avg ± SD**
		AYRSHIRES			
Milk (kg/day)	58746^1 24438^2 15491^3	8.3^1 8.2^2 8.1^3	1.1	54.6	21.42 ± 8.02
BF (kg/day)	97 363		0.03	2.68	0.82 ± 0.32
Prot (kg/day)	97 602		0.07	2.19	0.71 ± 0.25
SCS	22879^1 10561^2 7228^3	7.3^1 7.4^2 7.4^3	6.91	16.12	11.21 ± 1.17
		GUERNSEYS			
Milk (kg/day)	23690^1 11618^2 5762^3	7.8^1 7.6^2 7.5^3	1.4	63.5	16.41 ± 6.01
BF (kg/day)	41 070		0.06	2.63	0.72 ± 0.27
Prot (kg/day)	41 070		0.06	2.60	0.57 ± 0.20
SCS	10837^1 9375^2 6760^3	6.2^1 6.2^2 6.2^3	6.91	16.12	11.50 ± 1.15
		HOLSTEINS			
Milk (kg/day)	38945^1 28662^2 21548^3	5.9^1 5.9^2 5.9^3	2.5	77.8	22.82 ± 8.50
BF (kg/day)	87695		0.06	3.22	0.80 ± 0.31
Prot (kg/day)	87991		0.08	2.33	0.73 ± 0.26
SCS	8986^1 7511^2 5534^3	5.2^1 5.3^2 5.4^3	6.91	16.12	11.75 ± 1.31

		JERSEYS			
Milk (kg/day)	39183^1 15291^2 10724^3	7.4^1 7.5^2 7.5^3	1.2	44.2	15.27 ± 5.25
BF (kg/day)	64 114		0.02	2.66	0.69 ± 0.25
Prot (kg/day)	64126		0.04	1.63	0.56 ± 0.19
SCS	8782^1 4203^2 3656^3	5.9^1 5.5^2 5.4^3	6.91	15.90	11.56 ± 1.13

BF = butterfat yield
Prot = protein yield
SCS = Somatic Cell Score
SD = Standard Deviation
TD = Test-day records
$^{1\ 2\ 3}$ = Lactations 1, 2 and 3 respectively

Except for the Holstein evaluation, all multitrait analyses converged successfully. Several random selections were made on the Holstein data set and although convergence with status 1 (defined as standardized maximum change of the solutions from one round to the next with a stopping criterion of 0.001) , could not be obtained with any of the variance component estimation runs, estimates of the different runs were similar. Average estimates of the different runs are therefore presented without standard deviations for the Holsteins.

Variance component ratios and correlations for the different breeds are indicated in Table 5.3. Heritabilities for the animal effect of the Guernsey breed were higher for all traits compared to the other breeds, *i.e.* 24% for milk, 13% for butterfat and 19% for protein yield, as well as 8% for SCS. Estimates for the other breeds were all in the range of 17% for milk, 10.5% for butterfat, 15% for protein yield and 6% for SCS.

Table 5.3 Variance component ratios and correlations (± SE) for production traits and SCS.

		Milk	**BF**	**Prot**	**SCS**
			AYRSHIRES		
Animal	**Milk**	0.171 ± 0.009	0.891 ± 0.009	0.894 ± 0.007	
	BF		0.102 ± 0.006	0.917 ± 0.007	
	Prot			0.149 ± 0.008	
	SCS				0.064 ± 0.010
PE	**Milk**	0.245 ± 0.007	0.947 ± 0.004	0.980 ± 0.002	
	BF		0.175 ± 0.005	0.955 ± 0.003	
	Prot			0.225 ± 0.007	
	SCS				0.219 ± 0.009
Residual	**Milk**	0.585 ± 0.004	0.743 ± 0.001	0.938 ± 0.000	
	BF		0.723 ± 0.003	0.726 ± 0.001	
	Prot			0.626 ± 0.004	
	SCS				0.719 ± 0.005
			GUERNSEYS		
Animal	**Milk**	0.244 ± 0.023	0.852 ± 0.021	0.919 ± 0.010	
	BF		0.126 ± 0.017	0.895 ± 0.017	
	Prot			0.188 ± 0.021	
	SCS				0.081 ± 0.019
PE	**Milk**	0.314 ± 0.021	0.937 ± 0.007	0.978 ± 0.003	
	BF		0.251 ± 0.016	0.963 ± 0.005	
	Prot			0.306 ± 0.019	
	SCS				0.238 ± 0.016
Residual	**Milk**	0.443 ± 0.006	0.738 ± 0.002	0.915 ± 0.001	
	BF		0.623 ± 0.006	0.736 ± 0.002	
	Prot			0.505 ± 0.006	
	SCS				0.682 ± 0.007

			HOLSTEINS		
Animal	**Milk**	0.177	0.799	0.968	
	BF		0.105	0.866	
	Prot			0.136	
	SCS				0.062 ± 0.013
PE	**Milk**	0.382	0.879	0.848	
	BF		0.263	0.894	
	Prot			0.364	
	SCS				0.304 ± 0.013
Residual	**Milk**	0.441	0.662	0.926	
	BF		0.632	0.677	
	Prot			0.500	
	SCS				0.634 ± 0.006
			JERSEYS		
Animal	**Milk**	0.175 ± 0.013	0.779 ± 0.023	0.898 ± 0.010	
	BF		0.103 ± 0.009	0.882 ± 0.014	
	Prot			0.155 ± 0.012	
	SCS				0.063 ± 0.020
PE	**Milk**	0.304 ± 0.012	0.886 ± 0.008	0.952 ± 0.003	
	BF		0.193 ± 0.008	0.932 ± 0.006	
	Prot			0.259 ± 0.011	
	SCS				0.206 ± 0.018
Residual	**Milk**	0.521 ± 0.004	0.723 ± 0.001	0.918 ± 0.000	
	BF		0.703 ± 0.004	0.708 ± 0.001	
	Prot			0.586 ± 0.004	
	SCS				0.731 ± 0.007

BF = butterfat SCS = Somatic Cell Score PE = permanent environment

Danell (1982) found that heritability estimates of test-day yields were either the same or slightly lower than those of 305-day yields. Reents *et al.* (1995a) published heritability estimates averaging 26% for milk, 21% for butterfat and 23% for protein yields over 3 lactations, using fixed regression test-day models, treating the different lactations as different traits, while Lidauer *et al.* (2003) reported heritabilities of 26% for milk, 16% for protein and 17% for butterfat yields, respectively, for first lactation Finnish dairy cattle from a multitrait repeatability test-day model, *versus* 43% for milk, 29% for butterfat and 29% for protein yields on the same data set, using a multitrait lactation model. Pösö *et al.* (1996) also stated that heritability estimates for test-day milk yield tend to be generally low and distinctly lower than that of 305-day yields. Similar conclusions were made by Serrano *et al.* (2003) who also confirmed that low heritabilities were estimated with repeatability test-day animal models in sheep breeds. Heritability estimates obtained from this study were therefore on the lower range of test-day model estimates reported from the literature, but more comparable to these estimates than those obtained when fitting permanent environmental effects within lactation (Mostert *et al*., 2004b).

Reents *et al*. (1995b) concluded from various studies that the heritability for SCS ranges from 10-15% and in 1996 Mrode & Swanson published a weighted average of 0.11 ± 0.04, based on different measures of first-lactation SCS from several studies. However, Reents *et al.* (1995a) published heritabilities for SCS averaging 9.7% over three lactations for Canadian Holstein cows, using fixed regression test-day models, treating different lactations as different traits. De Haas *et al*. (2003) reported heritability estimates of 7-8% for lactational average somatic cell score for Dutch Friesian cows and Mrode & Swanson (2003) reported that heritabilities of individual test-days for SCS averaged 7% from a random regression test-day model. In contrast to these findings estimates as high as 19% (Pösö & Mäntysaari, 1996) and 23% (Luttinen & Juga, 1997) have been reported. Mrode & Swanson (2003) concluded that heritabilities from random regression models were higher than estimates from a repeatability model, using lactation averages regarding SCS. Heritablility estimates for SCS from this study were again on the lower range of estimates reported in the literature, but also more in agreement to these estimates than those obtained when fitting permanent environmental effects within lactation (Mostert *et al*., 2004a).

As can be expected from a repeatability model, permanent environmental effects accounted for a large proportion of the phenotypic variation, with the Holstein breed allocating the most for all traits and the Ayrshire breed the least for the production traits to permanent environment. For SCS the Jersey breed yielded the smallest estimate for permanent environmental effect of all the breeds. Permanent environmental variance of the production traits for the South African breeds were 1.25 (protein yield of the Guernsey breed) to 2.68 (protein yield for the Holstein breed) times higher than the genetic variance. For SCS permanent environmental variances were even higher, ranging from 2.95 (Guernsey breed) to 4.86 (Holstein breed) times the genetic variance. Da *et al*. (1992) found permanent environmental variance of SCS to be two times higher than the genetic variance for Holsteins. Reents *et al.* (1995a) published permanent environmental variances being on average over three lactations 1.6 times higher than genetic variance for production traits, and 4.3 times higher for SCS for Ontario Holstein cows, using fixed regression test-day models, while Serrano *et al.* (2003) reported permanent environmental variance to be four times higher than the genetic variance for SCS of Manchega ewes. According to Serrano *et al.* (2003) high permanent environmental variance relative to the genetic, indicates some degree of confounding, which may be due to the lack of pedigree information and insufficient repeated measures per animal. This might explain why permanent environmental variance for SCS of all the breeds is much higher than the genetic variance, since records for somatic cell count have only been captured on the INTERGIS since 1996, compared to 1988/1989 for the production traits, yielding more than a generation less historical data for SCS evaluations. Druet *et al.* (2003) and Mrode & Swanson (2003) also reported a large influence of permanent environmental effects, especially in the early stage of lactation, when using random regression test-day models.

Residual variance for production traits were highest in the Ayrshire breed and in the Jersey breed for SCS, with the Guernsey and Holstein analyses producing around 10% lower estimates for production traits. It is interesting to note that the residual variation associated with butterfat yield is much larger for all the breeds, compared to that of milk and protein yields. This is probably due to butterfat yield being more sensitive to nutritional practices (Pennington, 2005; Shirley *et al*., 1998; Socha & Johnson, 1998). Residual variances from random regression test-day models reported in the literature are smaller than the permanent environmental variance (Misztal *et al*., 2000; Druet *et al*., 2003 and De Roos *et al*., 2004), which is in contrast with this study, where the residual variance are greater than genetic and permanent environmental variances for all traits and all breeds. This, however, can be attributed to the modeling of genetic and non-genetic variance covariance structures between any pair of test-day observations along the course of the lactation by random regression models (Lidauer *et al*., 2003), whereas for test-day repeatability models the random genetic animal and permanent environmental effects are modeled as constants for each day in milk (Jensen, 2001). Because random regression models allow for a more comprehensive description of the animals' genotype (Lidauer *et al.,* 2003), associated residual effects should be smaller compared to that of fixed regression test-day models. Reents *et al*. (1995b) also reported residual variances that are greater than genetic and permanent environmental variances for somatic cell score, using fixed regression test-day models and treating different lactations as different traits.

Genetic correlations between traits were lowest between milk and butterfat yields for all breeds, ranging from 0.779 ± 0.023 for the Jersey breed to 0.891 ± 0.009 for the Ayrshire breed. Except for the Ayrshire breed which had a higher genetic correlation between butterfat and protein yield (0.917 ± 0.007), all breeds showed the highest correlation between milk and protein yields, ranging from 0.898 ± 0.010 for the Jersey to 0.968 for the Holstein. Not many references are available on genetic correlations between production traits based on test-day evaluations, as most test-day studies include different lactations as different traits and do evaluations separately for the yield traits. Pander *et al.* (1992) as well as Lidauer *et al*. (2003), however, also reported genetic correlations on different days in milk that is higher between milk and protein yield, compared to milk and butterfat yield on the same days in milk.

Genetic correlations of 0.80 between milk and protein yield, 0.60 between milk and fat yield and 0.70 between butterfat and protein yield from a multitrait repeatability test-day model, was indicated by Lidauer *et al.* (2003).

The correlations between permanent environmental effects showed more or less the same tendency as the genetic correlations, with highest correlations between milk and protein yields, ranging from 0.952 ± 0.003 for the Jersey breed to 0.980 ± 0.002 for the Ayrshire breed and lowest correlations between milk and butterfat yields, ranging from 0.886 ± 0.008 for the Jersey breed to 0.947 ± 0.004 for the Ayrshire breed. Although the permanent environmental ratios were highest for the Holstein breed, correlations between the traits for these effects were lower and also in a different order compared to the other breeds. For the Holstein breed the correlations between permanent environmental effects ranged from 0.848 between milk and protein yields, to 0.894 between butterfat and protein yields.

In Tables 5.4-5.7 genetic correlations between South Africa and other countries participating in INTERBULL (MACE evaluations) for the Ayrshire, Guernsey, Holstein and Jersey breeds, are presented (www-INTERBULL.slu.se). These correlations were estimated during the March 2005 INTERBULL test-runs.

A total of 86 Ayrshire sires (54 for SCS) from South Africa qualified for participation in the March 2005 INTERBULL test-run. South Africa contributed the least sires and Sweden the most sires (2508) to this test-run. Most common sires were shared with Canada and no sires with Denmark, Estonia or Norway. The lowest correlations for the production traits were obtained with New Zealand for butterfat and protein yields, as well as with Australia for butterfat yield (0.76), while the highest correlations were estimated for the countries having no common sires with South Africa (0.90 and 0.91). This is due to estimates being combined with those for Holstein where few or no common bulls are available for breeds other than Holstein. Genetic correlations between other countries ranged from 0.95 for milk yield between Estonia and Norway, sharing six common sires, to 0.76 between New Zealand and Canada, Finland, USA and Sweden for butterfat yield. Norway participate in the INTERBULL Udder Health evaluations with regards to clinical mastitis, but not for SCS. Genetic correlations with other countries were higher for SCS compared to those of the production traits.

Table 5.4 Number of common bulls for the production and (SCS) traits and genetic correlations between South Africa and other countries participating in MACE for the Ayrshire breed, as estimated by INTERBULL.

	Number of common bulls	Milk Yield	Butterfat Yield	Protein Yield	SCS
AUS	17 (10)	0.81	0.76	0.80	0.84
CAN	45 (32)	0.88	0.86	0.86	0.89
DEU	1 (1)	0.86	0.86	0.86	0.90
DNK	0 (0)	0.86	0.86	0.86	0.88
EST	0 (0)	0.90	0.90	0.90	0.92
FIN	4 (1)	0.89	0.89	0.89	0.90
GBR	12 (11)	0.86	0.86	0.86	0.90
NOR	0 (-)	0.90	0.90	0.91	-
NZL	18 (12)	0.77	0.76	0.76	0.80
SWE	11 (8)	0.87	0.87	0.87	0.91
USA	23 (15)	0.86	0.86	0.86	0.90

AUS – Australia
BEL – Belgium
CAN – Canada
CHE – Swiss Black & White
CHR – Swiss Red Holstein
CZE – Czech Republic
DEU – Germany
DNK – Denmark
DNR – Danish Red & White
ESP – Spain
EST – Estonia
FIN – Finland
FRA – France
FRR – French Pie Rouge
GBR – Great Brittan
HUN – Hungary
IRL – Ireland
ISR – Israel
ITA – Italy
JPN – Japan
NLD – Netherlands
NOR – Norway
NZL – New Zealand
POL – Poland
SVN – Slovenia
SWE – Sweden
USA – United States of America

From Table 5.5 it is clear that not many Guernsey bulls are shared between South Africa and other countries participating in MACE. The Guernsey breed is, worldwide, small in numbers, with the USA contributing the most AI-sires (576) and South Africa a minor 40 bulls to this test-run. Nevertheless, correlations between South Africa and these countries compared well with those between the rest of the participating countries, ranging from 0.76 between, amongst others, Canada and Australia for protein yield, to 0.95 between Canada and the USA for milk yield. With regards to the Guernsey breed, South Africa correlates least with New Zealand and Australia.

Table 5.5 Number of common bulls and genetic correlations between South Africa and other countries participating in MACE for the Guernsey breed, as estimated by INTERBULL.

	Number of common bulls	**Milk Yield**	**Butterfat Yield**	**Protein Yield**
AUS	6	0.78	0.77	0.77
CAN	2	0.86	0.86	0.86
GBR	4	0.86	0.86	0.86
NZL	4	0.77	0.77	0.77
USA	9	0.86	0.86	0.86

See Table 5.4 for the country codes.

Due to too few records being available for SCS of the Guernsey breed, South Africa does not yet participate in INTERBULL evaluations for this trait.

Table 5.6 Number of common bulls for the production and (SCS) traits and genetic correlations between South Africa and other countries participating in MACE for the Holstein breed, as estimated by INTERBULL.

	Number of common bulls	**Milk Yield**	**Butterfat Yield**	**Protein Yield**	**SCS**
AUS	164 (128)	0.76	0.75	0.75	0.91
BEL	84 (66)	0.85	0.85	0.85	0.94
CAN	204 (172)	0.85	0.85	0.85	0.91
CHE	87 (86)	0.86	0.85	0.85	0.95
CHR	32 (28)	0.86	0.86	0.86	0.92
CZE	84 (-)	0.86	0.86	0.86	-
DEU	201 (160)	0.85	0.85	0.85	0.90
DNK	104 (79)	0.85	0.85	0.85	0.88
DNR	2 (1)	0.90	0.90	0.91	0.91
ESP	190 (162)	0.85	0.85	0.85	0.97
EST	9 (8)	0.89	0.89	0.89	0.92
FIN	12 (12)	0.87	0.88	0.88	0.90
FRA	139 (121)	0.85	0.85	0.85	0.91
FRR	1 (1)	0.90	0.90	0.90	0.92
GBR	229 (185)	0.85	0.85	0.85	0.92
HUN	143 (114)	0.85	0.85	0.85	0.89
IRL	123 (-)	0.75	0.75	0.75	-
ISR	19 (13)	0.77	0.78	0.77	0.87
ITA	177 (146)	0.85	0.85	0.85	0.92
JPN	107 (94)	0.85	0.85	0.85	0.89
NLD	198 (149)	0.85	0.85	0.85	0.91
NZL	159 (132)	0.75	0.75	0.75	0.76

POL	63 (-)	0.86	0.86	0.86	-
SVN	20 (-)	0.87	0.87	0.87	-
SWE	95 (120)	0.85	0.85	0.85	0.92
USA	313 (255)	0.85	0.85	0.85	0.90

See Table 5.4 for the country codes.

The Holstein breed is by far the largest breed participating in INTERBULL. A total of 27 countries participated in the March 2005 test-run, with the USA contributing most sires (20 128) and France contributing least (112) sires to the evaluation. For South Africa 826 sires (616 for SCS) qualified for participation. With regards to common sires, French Pie Rouge (1) and Danish Red and Whites (2) shared the least and the USA the most (313) with South Africa. Correlations ranged from 0.75 with Australia (butterfat and protein yields), Ireland and New Zealand (all yield traits) to 0.91 with Danish Red and White for protein yield. Genetic correlations between other countries ranged from 0.95 between Canada and Japan, Denmark and Japan and the Netherlands and Japan for milk yield to 0.75 between amongst others, New Zealand and Canada for butterfat yield. A few countries do not yet participate in INTERBULL evaluations for SCS. Genetic correlations are generally higher for SCS in comparison with those for the production traits.

Germany contributed the least (47) and New Zealand the most sires (2 342) to the March 2005 Jersey test-run, with 431 sires (248 sires for SCS) qualifying for South Africa. The most common bulls were shared with the USA (136) and least with the Netherlands (20). Again the lowest correlations were obtained with New Zealand and Australia, 0.76 for butterfat and protein yields. With all other countries and for all the traits the genetic correlation with South Africa is 0.86, except with the USA, for which it is 0.85. Correlations between the other countries ranged from 0.95 between Canada and the USA for milk yield, to 0.74 between New Zealand and Great Britain for milk and butterfat yields. Germany and Italy do not participate in the INTERBULL Udder Health evaluations. Genetic correlations between South Africa and other countries are the same or higher for SCS compared to those for production traits.

Table 5.7 Number of common bulls for the production and (SCS) traits and genetic correlations between South Africa and other countries participating in MACE for the Jersey breed, as estimated by INTERBULL.

	Number of common bulls	**Milk Yield**	**Butterfat Yield**	**Protein Yield**	**SCS**
AUS	91 (65)	0.81	0.76	0.76	0.85
CAN	62 (49)	0.86	0.86	0.86	0.88
DEU	25 (-)	0.86	0.86	0.86	-
DNK	45 (34)	0.86	0.86	0.86	0.86
GBR	51 (41)	0.86	0.86	0.86	0.87
ITA	37 (-)	0.86	0.86	0.86	-
NLD	20 (13)	0.86	0.86	0.86	0.88
NZL	91 (72)	0.77	0.76	0.76	0.77
USA	136 (95)	0.85	0.85	0.85	0.89

See Table 5.4 for the country codes.

For all breeds genetic correlations between South Africa and other participating countries were in the mid-eighties to low nineties, except with New Zealand, Australia, Ireland and Israel. Although in general these correlations were lower than those amongst other countries, they were well within the range of 0.74 – 0.95, rendering compatibility with genetic methodologies of countries participating in international evaluations. The obvious lower correlations between New Zealand, Australia, Ireland and Israel with all other countries, are due to grouping of countries and setting of windows by INTERBULL so that estimates fall within certain windows (INTERBULL Technical Workshop, 2004, Uppsala, Sweden). For yield traits separate windows are maintained depending on the climate and whether or not countries predominantly have grazing systems. Two countries with similar climate and production system (grazing *vs.* non-grazing) are expected to be more correlated with each other than two countries with different climate or production systems. If estimates are higher than the maximum (or lower than the minimum) value, they are set to the maximum (or minimum) value. In addition, estimates are regressed towards a mean correlation within groups, the regression depending on the number of common bulls. Countries grouped together are therefore Australia, Ireland and New Zealand (grazing), Israel (climate) and all other countries. Windows used are correlation between countries belonging to the same group (>0.80), and correlation between countries belonging to different groups (>0.75). Furthermore, for breeds other than Holstein, estimates are combined with genetic correlations for Holstein, in order to obtain correlations between countries sharing few or no common bulls. Genetic correlations between countries are higher for udder health traits compared to those for production traits, due to different windows being used (Correlations between countries belonging to the same group : 0.85-0.98; Correlations between countries belonging to different groups : 0.75-0.98). The same grouping of countries is used for production and somatic cell score traits (INTERBULL Technical Workshop, 2004, Uppsala, Sweden).

5.5 CONCLUSIONS

Heritability estimates were obtained for production traits and somatic cell score that were on the lower range of literature estimates for test-day models, using a multi-lactation repeatability test-day model where the permanent environmental effect was fitted across lactations. These estimates were, however, more in agreement with literature estimates compared to estimates from the test-day models fitting the permanent environmental effect within lactation. Trend validation tests were accepted by INTERBULL during the March 2005 test-runs for all the breeds and traits, except for SCS of the Guernsey breed, due to too few records available for this breed. Genetic correlations between South Africa and other participating countries, estimated by INTERBULL during the March 2005 test-runs, were around 0.80, which is close to those amongst the other participating countries. MACE (multiple across country evaluation) breeding values will therefore be obtained after the May 2005 INTERBULL routine run, implementing the new methodology for South Africa's Ayrshire, Guernsey, Jersey and Holstein breeds.

5.6 REFERENCES

Da, Y., Grossman, M., Misztal, I. & Wiggans, G.R., 1992. Estimation of genetic parameters for somatic cell score in Holsteins. J. Dairy Sci. 75, 2265-2271.

Danell, B., 1982. Studies on lactation yield and individual test-day yields of Swedish dairy cows. II. Estimates of genetic and phenotypic parameters. Acta Agric. Scand. 32, 82-92.

De Haas, Y., Barkema, H.W., Schukken, Y.H. & Veerkamp, R.F., 2003. Genetic associations for pathogen-specific clinical mastitis and patterns of peaks in somatic cell count. Anim. Sci. 77, 187-195.

De Roos, A.P.W., Harbers, A.G.F. & De Jong, G., 2004. Random herd curves in a test-day model for milk, fat, and protein production of dairy cattle in The Netherlands. J. Dairy Sci. 87, 2693-2701.

Druet, T., Jaffrézic, F., Boichard, D. & Ducrocq, V., 2003. Modeling lactation curves and estimation of genetic parameters for first lactation test-day records of French Holstein cows. J.Dairy Sci. 86, 2480-2490.

Groeneveld, E. & Garcia-Cortes, A., 1998. Proc. 6th WCGALP. 27, 455-456.

Groeneveld, E. & Kovac, M., 1990. A generalized computing procedure for setting up and solving mixed linear models. J. Dairy Sci. 73, 513-531.

Haile-Mariam, M., Goddard, M.E. & Bowman, P.J., 2000. Estimates of Genetic Parameters for Daily Somatic Cell Count of Australian Dairy Cattle. J.Dairy Sci. 84, 1255-1264.

INTERBULL Technical Workshop, 2004. Genetic correlation estimation procedure. January 2004. Uppsala, Sweden.

Jamrozik, J., Schaeffer, L.R. & Dekkers, J.C.M., 1997. Genetic evaluation of dairy cattle using test-day yield and random regression model. J. Dairy Sci. 80, 1217-1226.

Jensen, J., 2001. Genetic evaluation of dairy cattle using test-day models. J. Dairy Sci. 84, 2803-2812.

Klei, B. & Weigel, K., 1998. A method to estimate correlations among countries using data on all bulls. Proc. 1998 INTERBULL Meeting, Rotorua, New Zealand, Jan 18-19, 8-14.

Lidauer, M., Mäntysaari, E.A. & Strandén, I., 2003. Comparison of test-day models for genetic evaluation of production traits in dairy cattle. Livest. Prod. Sci. 79, 73-86.

Luttinen, A. & Juga, J., 1997. Genetic relationships between milk yield, somatic cell count, mastitis and leakage in Finnish dairy cattle population. In: Proceedings of the International Workshop on Genetic Improvement of Functional Traits in Cattle: Health. Uppsala, Sweden, June, INTERBULL Bull. No. 15, 78-83.

Misztal, I., Strabel, T., Jamrozik, J., Mäntysaari, E. A. & Meuwissen, T.H.E., 2000. Strategies for estimating the parameters needed for different test-day models. J. Dairy. Sci. 83, 1125-1134.

Mostert, B.E., Banga, C.B., Groeneveld, E. & Kanfer, F.H.J., 2004a. Breeding value estimation for somatic cell score in South African dairy cattle. S. Afr. J. Anim. Sci. 34, 32-34.

Mostert, B.E., Groeneveld, E. & Kanfer, F.H.J., 2004b. Testday models for production traits in dairy cattle. S. Afr. J. Anim. Sci. 34, 35-37.

Mostert, B.E., Theron, H.E. & Kanfer, F.H.J., 2001. The effect of calving season and age at calving on production traits of South African dairy cattle. S. Afr. J. Anim. Sci. 31, 205-214.

Mrode, R.A. & Swanson, G.J.T., 1996. Genetic and statistical properties of somatic cell count and its suitability as an indirect means of reducing the incidence of mastitis in dairy cattle. Anim. Breed. Abstr. 64, 847-857.

Mrode, R.A. & Swanson, G.J.T., 2003. Estimation of genetic parameters for somatic cell count in the first three lactations using random regression. Livest. Prod. Sci. 79, 239-247.

Pander, B.L., Hill, W.G. & Thompson, R., 1992. Genetic parameters of test-day records of British Holstein-Friesian heifers. Anim. Prod. 55, 11-21.

Pennington, J.A., 2005. Factors affecting fat percent in milk of lactating cows. Agriculture and Natural Resources. University of Arkansas Cooperative Extension Service, Little Rock.

Pösö, J. & Mäntysaari, E.A., 1996. Relationships between clinical mastitis, somatic cell score and production for the first three lactations in Finnish Ayrshires. J. Dairy Sci. 79, 1284-1291.

Pösö, J., Mäntysaari, E.A. & Kettunen, A., 1996. Estimation of genetic parameters of test-day production in Finnish Ayrshire cows. Proc. 1996 INTERBULL Meeting, Veldhoven, The Netherlands, June 23-24, 45-48.

Reents, R., Dekkers, J.C.M. & Schaeffer, L.R., 1995b. Genetic evaluation for somatic cell score with a test-day model for multiple lactations. J. Dairy Sci. 78, 2858-2870.

Reents, R., Jamrozik, J., Schaeffer, L.R. & Dekkers, J.C.M., 1995a. Estimation of genetic parameters for test-day records of somatic cell score. J. Dairy Sci. 78, 2847-2857.

Serrano, M., Pérez-Guzmán, M.D., Montoro, V. & Jurado, J.J., 2003. Genetic analysis of somatic cell count and milk traits in Manchega ewes. Mean lactation and test-day approaches. Livest. Prod. Sci. 84, 1-10.

Shirley, J.E., Piehl, D. & Titgemeyer, E.C., 1998. Effect of level of dietary protein on the performance of lactating dairy cows. 1998 Joint Meeting Abstracts. J. Dairy Sci. 81, Suppli. 1, J. Anim. Sci. 76 Suppl.1, 251.

Socha, T. & Johnson, A.B., 1998. Summary of trials conducted evaluating the effect of a combination of complex zinc methionine, manganese methionine, copper lysine and cobalt glucoheptonate on lactation and reproduction performance of dairy cattle. 1998 Joint Meeting Abstracts. J. Dairy Sci. 81, Suppl. 1; J. Anim. Sci. 76, Suppl.1, 251.

Swalve, H.H., 1998. Use of test-day records in genetic evaluation. Proc. 6th World Congr. Genet. Appl. Livest. Prod., Armidale, New South Wales, Australia. 23, 295-302.

Swalve, H.H., 2000. Symposium: Test-Day Models. Theoretical Basis and Computational Methods for Different Test-Day Genetic Evaluation Methods. J. Dairy Sci. 83, 1115 – 1124.

Wilmink, J.B.M., 1987. Adjustment of test-day milk, fat and protein yield for age season and stage of lactation. Livest. Prod. Sci. 16, 335-348.

CHAPTER 6

COMPARISON OF BREEDING VALUES AND GENETIC TRENDS FOR PRODUCTION TRAITS ESTIMATED BY A LACTATION MODEL AND A FIXED REGRESSION TEST-DAY MODEL

B.E.Mostert[#], H.E. Theron, F.H.J. Kanfer[1] & E. van Marle-Köster[2]
ARC-LBD, Private Bag X2, Irene, 0062, South Africa
[1]Department of Statistics, University of Pretoria, Pretoria, 0002, South Africa
[2]Department of Animal and Wildlife Sciences, University of Pretoria, 0002, Pretoria, South Africa

---

6.1 ABSTRACT

A comparison of breeding values and genetic trends of production traits, estimated by the September/October 2003 South African National Genetic Evaluation, using a Lactation Model (LM) and the 2004 South African National Genetic Evaluation, using a Fixed Regression Test-day Model (TDM) is made for Ayrshire, Guernsey, Holstein and Jersey cows participating in the South African Dairy Animal Improvement Scheme. Specific differences between the two models are documented, with differences in statistical methodology and inclusion of test-day records of the first three parities in the TDM versus only first lactation 305-day yields in the LM, as the main differences. Significant reranking of especially cows and young sires occurred between the models. Genetic trends of the TDM were not as steep as those from the LM, as the trait that was selected for was first lactation yield, while the TDM trends reflect genetic progress over the first three parities.

Keywords : Ayrshire, correlations, dairy cattle, Guernsey, Holstein, Jersey, ranking.
[#]Corresponding author. E-mail : Bernice@arc.agric.za

6.2 INTRODUCTION

Genetic evaluation of production traits in dairy cattle can be improved by using test-day observations (Test-day Models) instead of aggregated 305-day production records (Lactation Models) (Ptak & Shaeffer, 1993; Reents *et al*., 1995a; Jamrozik *et al*., 1997; Schaeffer *et al*., 2000 & Lidauer *et al*., 2003). Breeding values, based on production in the first three lactations and estimated with Fixed Regression Multiple Lactation Repeatability Test-day Models, have been released in South Africa during 2004 for Ayrshire, Guernsey, Jersey and Holstein breeds. As in other countries, changing from a Lactation to a Test-Day Model caused significant changes in the ranking of animals. After the Canadian Test-Day Model evolved over a period of seven years, Schaeffer *et al.* (2000) remarked that efforts to inform the dairy industry about changes in EBV due to the new model, were significantly challenging. This can definitely be confirmed by the authors. The aim of this paper is to document the differences between the two models, to quantify the rank changes in the South African dairy populations and to make recommendations for improvement of the Test-day Model for future reference.

6.3 MATERIALS AND METHODS

Milk, butterfat and protein breeding values (EBVs) from the September/October 2003 South African National Genetic Evaluation were based on first lactation 305-day production records and estimated using a Lactation Model (LM). These EBVs were compared to milk, butterfat and protein breeding values estimated in the 2004 South African National Genetic Evaluation, based on the first three lactations using test-day records, estimated with a Fixed Regression Multiple Lactation Repeatability Test-day Model (TDM). This was done for Ayrshire, Guernsey, Holstein and Jersey cows participating in the South African Dairy Animal Improvement Scheme (Mostert *et al*., 2006). A TDM, using only first lactation test-day records, were also run for the Holstein and Guernsey breeds to compare breeding values from first lactation test-day records, with breeding values from first lactation 305-day records (LM), as well as with breeding values from the TDM, including test-days from the first three lactations. Pearson correlations were obtained using SAS (1996) for animals which had breeding values in both evaluations. Bulls were grouped into Proven Sires, having daughters in at least 10 herds, and Unproven Sires, having daughters in less than 10 herds. Only measured cows were compared. Genetic trends for milk yield from the TDM and LM were obtained by averaging the breeding values of measured cows per birth year (TDM), or lactation year (LM).

6.4 RESULTS AND DISCUSSION

Records of second and later lactations provide more complete information on lifetime performance than using records from first lactations only (Powell & Norman, 1981). The inclusion of records of the first three lactations in the TDM, was therefore a welcomed and long awaited step forward for the dairy industry. It had, however, a significant effect on the ranking of animals when compared to breeding values of the LM, which were based on first lactation records only. This particularly influenced “late bloomers” and their sires, only producing to their full potential in the second and later lactations and now being recognized by the TDM. Also, the inclusion of short lactations (lactations ending before 240 days in milk) in the TDM allows for penalization of sires breeding a high percentage of daughters having short lactations. In the LM, these records were not included. As short lactations are mainly due to selection against low production, certain sires’ breeding values were therefore previously overestimated with the LM. A part of ranking differences between the two models is of course due to inclusion of new data in the TDM since the 2003 LM evaluations, for all breeds.

A disadvantage of implementing the TDM, however, is that lactation records have been stored on the INTERGIS (Integrated Registration and Genetic Information System) since 1978, but test-day records only since 1988/89, resulting in the TDM being based on two generations less data, compared to the LM. This has a particular effect on the accuracy of evaluations of dams of currently active cows, as their breeding values are now based solely on progeny and not their own performance. It was also found that older sires, especially those born during the 1980s, who performed well at their time, but whose dams’ and other female relatives’ test-day records were not stored on the INTERGIS and who did not have active progeny in the TDM anymore, ranked too high. This is probably due to BLUP not being able to find links to correctly rank these sires in the population. These effects will gradually disappear after a sufficient number of years of test-day records have accumulated. The INTERBULL Guidelines for National and International Evaluation Systems (2001) recommends that genetic evaluation for production traits should be based on at least 15 years of data. For the production evaluations using the TDM, this recommendation is met, but for somatic cell score, where records have been stored on the INTERGIS only since 1996, a few years of data are still lacking to meet this criterion.

The TDM allows for records from cows with unknown parents, while the LM required that at least one parent should be known. Also, records where only milk yield measurements are available were included in the TDM, while milk, butterfat and protein yields should have been known for a cow to be included in the LM.

Genetic grouping in the pedigree is important in achieving accurate evaluations by defining different populations from which animals arise (Pollak & Quaas, 1983). Unknown parents in the pedigree are replaced by genetic groups, which provide a substitute genetic value for an unknown parent. The genetic value of such a group is determined by all animals, which are connected to this specific group (Handbook NRS, 2004). The INTERBULL Guide for National and International Evaluation Systems (2001) recommends that genetic groups should have a minimum of 10-20 animals. Some of these groups consisted of less than 10 animals in the LM. It was therefore decided to redefine the genetic groups for all breeds when the TDM was implemented, e.g. for the Jersey breed 37 groups were defined in the pedigree for the TDM *versus* 90 groups in the LM. As the function of genetic groups is specific for each animal and depends on the number of generations to the base phantom ancestors and on the genetic groups to which those phantom ancestors are assigned (Westell *et al*., 1988), changes in genetic group definition caused differences in the breeding values of the two models. Redefining of these genetic groups caused the breeding values from the TDM to follow a narrower spread for all breeds, as the evaluations now accommodate less groups in the pedigree compared to the LM evaluations.

The modeling of fixed and random effects for breeding value estimation differed between the LM and TDM, as completely different methodologies and data structures were being used. For the LM the contemporary group was defined as herd x test year x status (commercial *versus* registered cows), age at calving was included as a fixed effect (6 classes) and a sire x herd interaction was also fitted as a random effect to account for possible preferential treatment. For the TDM the contemporary group was defined as herd x test-date x parity x number of milkings. This is probably the most important advantage of the TDM compared to the LM, as factors specific to each test-day can be accounted for (Jamrozik *et al*., 1997). Meyer *et al.* (1989) and Rekaya *et al.* (1995) reported that fitting a herd x test-date effect reduced the residual variances in comparison to analyses comparing tests in the same herd x year x season, while Swalve (1995) found an increase in additive and permanent environmental variance and a decrease in residual variance for herd x test-date models. Ptak & Schaeffer (1993) concluded that with the herd x test-date effect, a more precise adjustment for seasons as they affect each test-day, can be made. In the TDM age at calving was included as a fixed effect (14 classes across lactations) to account for the effect of age at calving on yield, the Wilmink curve was modeled on days in milk within lactation (3 classes) and season (2 classes) to account for the effect of stage of lactation on yield, calving interval was included as a fixed effect (9 classes) to account for the effect of length of the calving interval on milk yield in the following lactation and a permanent environmental effect (random) was fitted across lactations to account for a common effect of the environment associated with all test-day records of a cow. Variance components estimated with these models were considerably lower for the TDM, compared to the LM for all the breeds. It should, however, be noted that different traits are involved in these evaluations, *i.e.* production per day, analyzed by the TDM and accumulated 305-day production, analyzed by the LM. Lower heritabilities from the TDM are however, compensated for by the larger number of observations : up to ten test-day records per lactation and up to three lactations in the TDM *versus* one first lactation observation in the LM. Furthermore, because short-term environmental effects are modeled through the herd x test-date effect, it is expected that reliabilities of breeding values from the TDM will be higher compared to that of the LM. Consequently the spread of breeding values is expected to be larger for the TDM compared to the LM (Dr Freddy Fikse, 2006: Personal communication, INTERBULL Centre Director, e-mail: Freddy.Fikse@hgen.slu.se). As mentioned earlier that was, however, not the case due to much less genetic groups defined in the pedigrees of the TDM compared to the LM, influencing the overall spread of the breeding values of the populations.

Both models were multitrait evaluations, where milk, butterfat and protein yields were analyzed together in order to utilize the genetic correlations between the traits for accurately estimating the breeding values. This is especially important for future evaluations, as, since 2001, the South African Dairy Animal Improvement Scheme allows breeders to only record milk yields with measurement of butterfat and protein yields becoming optional (Dr. J. van der Westhuizen, 2001: Personal communication, INTERGIS Manager, e-mail: japie@arc.agric.za).

As the INTERBULL Guidelines for National and International Evaluation Systems (2001) recommends that the base year should change in 2005 to 2000, it was decided to implement this change already in 2004. A second adaptation by the industry following the introduction of breeding values from the TDM, was thereby avoided. Because three lactations were included in the TDM evaluations, the base year had to be redefined from cows tested in 1995, to cows born in 2000. This resulted, therefore, in a drastic change (around 7 years) in the base year, with breeding values of the whole population moving down. This is, however, only a scale effect and the ranking of the animals was not influenced. For the Ayrshire and Guernsey breeds, being small populations, this base change was too drastic, resulting in few animals having positive breeding values from the TDM. It was decided then to use cows born in 1995 as their base, a shift of around three years.

In Figures 6.1-6.4 genetic trends for milk yield for the Ayrshire, Guernsey, Holstein and Jersey breeds are indicated. Genetic trends are expressed on year of birth for breeding values of the TDM and not test year, as was done with breeding values from the LM. Since 1987 when breeding values were first introduced to the industry using a Sire Model (Loubser *et al*., 2001), the trait that has been selected for, was first lactation yield. Trends from the TDM are therefore not as steep as those from the LM, as they indicate genetic progress over the first three lactations, while trends from the LM indicate progress in first lactation yield, which was the selected trait. Furthermore, inclusion of short lactations in the TDM will also have a declining effect on the trends. The difference in the base year can also be observed from the graphs : cows born in 2000 (TDM) *versus* cows tested in 1995 (LM) for the Holstein and Jersey breeds and cows born in 1995 (TDM) *versus* cows tested in 1995 (LM) for the Ayrshire and Guernsey breeds, causing a time lag between the trends.

Figure 6.1 Genetic trends for milk yield of Ayrshire cattle.

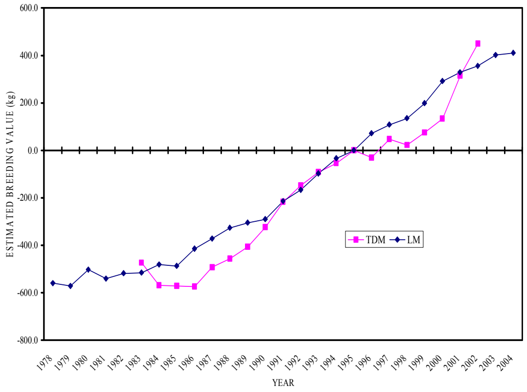


Figure 6.2 Genetic trends for milk yield of Guernsey cattle.

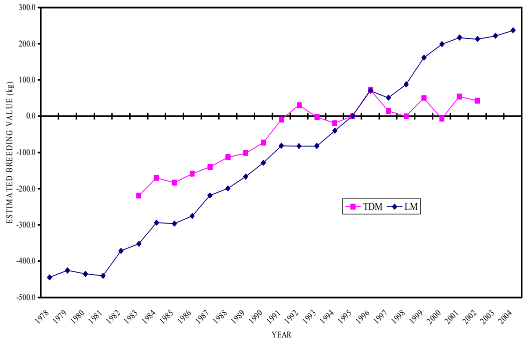


Figure 6.3 Genetic trends for milk yield of Holstein cattle.

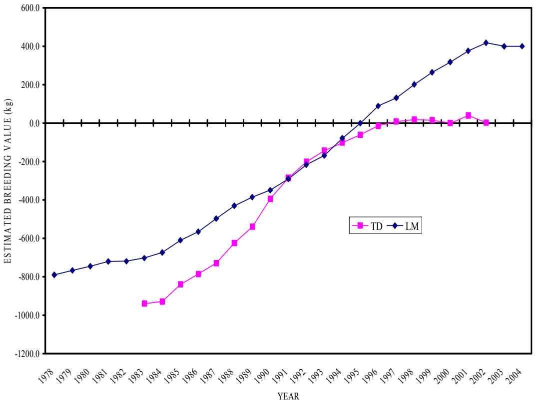
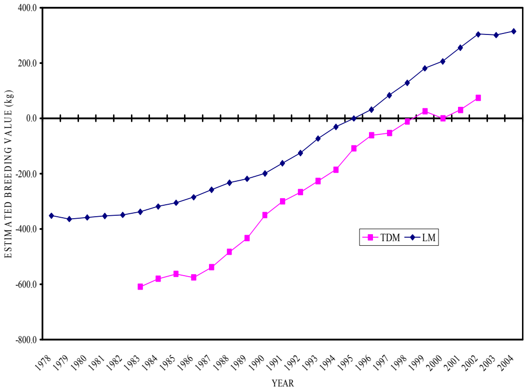


Figure 6.4 Genetic trends for milk yield of Jersey cattle.



Trends from the LM show clearly that significant progress was achieved in first lactation milk yield since 1978 in all breeds. Linear regressions quantified this progress as 41.9kg/year ($R^2 = 97\%$) for the Ayrshires, 29.0kg/year for the Guernseys ($R^2=99\%$), 53.6kg/year ($R^2 = 98\%$) for the Holsteins and 29.5kg/year for the Jerseys ($R^2=98\%$). Progress in milk yield over the first three lactations (TDM-trends), showed, however, different pictures for the breeds. For the Ayrshires almost the same progress was achieved in yield over the first three lactations compared to yield in the first lactation, with a slightly higher regression coefficient of 50.1kg/year ($R^2 = 94\%$). Initially progress in the first three lactations of the Guernsey breed was almost parallel to progress in the first lactation, but since 1992 the trend does not change significantly. Progress in milk yield of the first three lactations was 13.8kg/year ($R^2 = 82\%$) for the Guernsey breed. Although the regression coefficient for the TDM-trend of the Holsteins was higher (58.1kg/year; $R^2 = 90\%$) compared to that of the LM-trend, progress was achieved in the earlier years up to 1996, after which it stabilized. The Jersey breed, however, also showed comparable progress in the first three lactations with that of the LM-trend, with a regression coefficient of 40.2kg/year ($R^2 = 98\%$). In fact, the difference between the two trends for the Jersey breed can largely be attributed to the difference in the base year adaptation. It is, however, important to note that all the data for the last year of the trends were not yet included in the evaluations and that the different heritabilities involved in the TDM and LM also influence the rate of genetic progress. Reents *et al.* (1998) showed that the genetic trend is significantly higher for test-day models in comparison to lactation models. The same number of lactations was, however, involved in their study.

In Table 6.1 Pearson Correlations between breeding values from the Test-day Model and Lactation Models are indicated for proven sires, unproven sires and measured cows for the Holstein and Guernsey breeds and in Table 2 for the Ayrshire and Jersey breeds. Ranks between breeding values of the TDM and LM (Table 6.1 and 6.2) are in the high eighties and low nineties for proven sires of all breeds, with milk yield ranking the highest of the yield traits. For the small breeds, rankings were lowest for unproven sires for all traits compared to those of proven sires and cows. For the Holstein breed unproven sires ranked lowest for all traits and for the Jersey breed unproven sires and cows ranked the same, except for butterfat yield where the unproven sires ranked lower than the cows.

Table 6.1 Pearson Correlations between breeding values from three models for proven, unproven and measured cows of the Holstein and Guernsey breeds.

	HOLSTEIN			**GUERNSEY**		
	Milk	**Butterfat**	**Protein**	**Milk**	**Butterfat**	**Protein**
Proven Sires						
TDM *vs.* LM	0.91	0.88	0.89	0.86	0.83	0.83
1st TDM *vs.* LM	0.94	0.91	0.92	0.89	0.87	0.87
1st TDM *vs.* TDM	0.96	0.95	0.96	0.90	0.91	0.89
Young Sires						
TDM *vs.* LM	0.89	0.87	0.87	0.63	0.61	0.62
1st TDM *vs.* LM	0.91	0.88	0.89	0.82	0.77	0.79
1st TDM *vs.* TDM	0.95	0.95	0.94	0.88	0.85	0.86
Measured Cows						
TDM *vs.* LM	0.88	0.84	0.86	0.72	0.67	0.69
1st TDM *vs.* LM	0.92	0.87	0.91	0.83	0.81	0.80
1st TDM *vs.* TDM	0.93	0.94	0.92	0.85	0.82	0.85

1st TDM – Test-day Model including only test-day records of the first lactation
TDM – Test-day Model including test-day records of the first three lactations
LM – Lactation Model including 305-day yields of first lactations

These rankings were within the ranges reported in the literature. Reents *et al.* (1995b) reported correlations between breeding values for protein yield from a fixed regression test-day model and breeding values from a lactation average model, including test-days from the first three lactations, of 84% for sires having more than 100 daughters and 73-81% for sires having less than 100 daughters. Correlations between cow breeding values were in the range similar to correlations for bulls with a low number of daughters (*i.e.* 78%). In 1998 Reents *et al.* concluded that ranking of bull breeding values from test-day models compared to a lactation model changed substantially and published correlations in the range of 94% and for cows in the range of 87%.

According to Schaeffer *et al.* (2000) rankings for Canadian bulls ranged from 88% for the Brown Swiss (one of the smallest breeds) to 97% for the Holstein (largest breed) when comparing breeding values from a LM with that of a random regression test-day model. Lidauer *et al.* (2003) published high correlations between active bull's breeding values (0.99 for milk, 0.98 for butterfat and protein) from a LM and random regression test-day models, but the animals and data of their study were carefully selected to be the same across different models in contrast to the other mentioned publications. However, for young cows the corresponding correlations were about 88%, which led to significant re-ranking of cows. Other rankings reported in the literature were all in the range of 87%-97% (Swalve, 1995; Reents & Dopp, 1996; Jamrozik *et al*., 1997; Lidauer *et al*., 2000 and Emmerling, *et al*., 2002).

Table 6.2 Pearson Correlations between breeding values of two models for proven sires, unproven sires and measured cows of the Ayrshire and Jersey breeds.

	AYRSHIRE			**JERSEY**		
	Milk	**Butterfat**	**Protein**	**Milk**	**Butterfat**	**Protein**
			Proven Sires			
TDM *vs.* LM	0.89	0.87	0.86	0.92	0.90	0.89
			Young Sires			
TDM *vs.* LM	0.80	0.81	0.80	0.89	0.87	0.88
			Measured Cows			
TDM *vs.* LM	0.86	0.86	0.85	0.89	0.88	0.88

TDM – Test-day Model including test-day records of the first three lactations
LM – Lactation Model including 305-day yields of first lactations

The impact of using the TDM compared to the LM (Table 1), can be seen from the correlation estimates of breeding values from the TDM, including only 1st lactation records and breeding values from the LM. Differences in methodology therefore cause proven sires of the Holstein breed to re-rank with 6-9% and proven Guernsey sires with 11-13%, depending on the trait. Inclusion of 2nd and 3rd parity records in the TDM cause rankings for proven sires to be 3% lower for all traits of the Holstein breed (differences between rankings of TDM and LM with 1st TDM and LM). For the Guernsey breed the rankings of the proven sires were 3% lower for milk yield, and 4% for butterfat and protein yields. The same trends can be observed for the unproven sires and cows of both breeds. Reents *et al.* (1995b) reported correlations of 3.6 – 7.3% lower when comparing correlations between protein breeding values from a fixed regression test-day model and breeding values from a lactation average model, including test-days from the first three lactations, with correlations from a fixed regression test-day model, including test-day records from the first three lactations and breeding values from a lactation average model, including test-days from only the first lactation, depending on the number of daughters involved.

When using the same methodology, the effect of inclusion of 2nd and 3rd lactation records (Table 1) on rankings can be seen from the correlation estimates between breeding values based on a TDM including only first lactation records and a TDM including records from the first three lactations. This caused re-rankings of 4-5% in proven Holstein sires and 9-11% in proven Guernsey sires, depending on the trait. Again the impact of using different methodologies can be quantified when comparing these correlation estimates with those obtained between breeding values for a TDM including only 1st lactation records and breeding values from the LM. This cause the ranks to decrease with 2-4% for proven sires of the Holstein breed and 1-4% for proven sires of the Guernsey breed, depending on the trait. The same trends can be observed for the unproven sires and cows of both breeds.

Differences in methodology therefore had a slightly larger effect on the rankings of animals than inclusion of 2nd and 3rd lactation records.

6. 5 CONCLUSIONS

Implementation of the Fixed Regression Multiple Lactation Repeatability Test-day Model to estimate breeding values for production traits of dairy breeds caused significant re-rankings of especially cows and young sires. Genetic trends now reflect genetic progress in production over the first three lactations and not only first lactation progress, as with the previously used LM.

6.6 REFERENCES

Emmerling, R., Lidauer, M., Mäntysaari, E.A., 2002. Multiple lactation random regression test-day model for Simmental and Brown Swiss in Germany and Austria. Interbull Meeting, Interlaken Switzerland, 26-27 May, 111-117.

Handbook NRS, 2004. Statistical Indicators. E-7. Breeding value estimation of milk production traits with test-day model. CR Delta.

INTERBULL Guidelines for National and International Genetic Evaluation Systems in Dairy Cattle with Focus on Production Traits, 2001. Bulletin No. 28. Department of Animals Breeding and Genetics, Uppsala, Sweden.

Jamzorik, J., Schaeffer. L.R. & Dekkers, J.C.M., 1997. Genetic evaluation of dairy cattle using test-day yield and random regression model. J. Dairy Sci. 80, 2550-2556.

Lidauer, M., Mäntysaari, E.A. & Strandén, I., 2003. Comparison of test-day models for genetic evaluation of production traits in dairy cattle. Livest. Prod. Sci. 79, 73-86.

Loubser, L.F.B, Banga, C.B., Scholtz, M.M. & Hallowell, G.J. (Eds.), 2001. Dairy Herd Improvement in South Africa. The use of breeding values in the South African Dairy Industry. 25.

Meyer, K., Grase, H.-U. & Hammond, K., 1989. Estimation of genetic parameters for first lactation test-day production of Australian Black and White cows. Livest. Prod. Sci. 21, 177-199.

Mostert, B.E., Theron, H.E., Kanfer, F.H.J. & Van Marle-Köster, E., 2006. Test-day models for South African dairy cattle for participation in international evaluations. S. Afr. J. Anim Sci, 36, 58-70.

Pollak, E.J. & Quaas, R.L., 1983. Definition of group effects in sire evaluation models. J. Dairy Sci. 66, 1503-1509.

Powel, R.L. & Norman, H.D., 1981. Different lactations for estimating genetic merit of dairy cows. J.Dairy Sci. 64, 321-330.

Ptak, E., Schaeffer, L.R., 1993. Use of test-day yields for genetic evaluation of dairy sires and cows. Livest. Prod. Sci. 34, 23-34.

Reents, R., Dekkers, J.C.M. & Schaeffer, L.R., 1995a. Genetic evaluation for somatic cell score with a test-day model for multiple lactations. J. Dairy Sci. 78, 2858-2870.

Reents, R., Dekkers, J.C.M. & Schaeffer, L.R., 1995b. Genetic evaluation with a multiple lactation test-day model for SCS and production traits. Interbull Meeting, Prague, Czech Republic, September 7-8, 1-6.

Reents, R. & Dopp, L., 1996. Genetic evaluation for dairy production traits with a test-day model for multiple lactations. Interbull Meeting, Veldhoven, The Netherlands, 23-24 June, 113-117.

Reents, R., Dopp, L., Schmutz, M. & Reinhardt, F., 1998. Impact of application of a test-day model to dairy production traits on genetic evaluations of cows. Interbull Meeting, Rotorua, New Zealand, January 18-19, 49-54.

Rekaya, R., Béjar, F., Carabano, M.J. & Alenda, R., 1995. Genetic parameters for test-day measurements in Spanish Holstein-Friesian. Interbull Meeting, Prague, Czech Republic, September, 1-8.

SAS, 1996. SAS user's guide : Statistics, Release 6.12. SAS Institute Inc., Cary, North Carolina, USA.

Schaeffer, L.R., Jamrozik, J., Kistemaker, G.J. & Van Doornmaal, B.J., 2000. Experience with a test-day model. J. Dairy Sci. 83, 1135-1144.

Swalve, H.H., 1995. The effect of test-day models on the estimation of genetic parameters and breeding values for dairy yield traits. J. Dairy Sci. 78, 929-938.

Westell, R.A., Quaas, R.L. & Van Vleck, L.D., 1988. Genetic groups in an animal model. J. Dairy Sci. 71, 1310-1318.

CHAPTER 7

THE IMPACT OF THE ADJUSTMENT OF HETEROGENEOUS VARIANCES AND A CALVING YEAR EFFECT ON A TEST-DAY MODEL FOR NATIONAL GENETIC EVALUATION OF DAIRY CATTLE IN SOUTH AFRICA

B.E. Mostert#, H.E. Theron, F.H.J. Kanfer[1] and E. van Marle-Köster[2]
ARC-LBD, Private Bag X2, Irene, 0062, South Africa
[1]Department of Statistics, University of Pretoria, Pretoria, 0002
[2]Department of Animal and Wildlife Sciences, University of Pretoria, Pretoria, 0002

7.1 ABSTRACT

South Africa implemented test-day models for genetic evaluations of production traits, using a Fixed Regression Test-Day Model (FRTM), which assumes equal variances of the response variable at different days in milk, the explanatory variable. Data at the beginning and at the end of lactation period, have higher variances than tests in the middle of the lactation. Furthermore, first lactations have lower mean and variances compared to second and third lactations. This is a deviation from the basic assumptions required for the application of repeatability models. A modification was therefore implemented to reduce the effect of deviating from this assumption. Test-day milk, butterfat and protein yield records of Jersey cows, participating in the South African Milk Recording Scheme, were therefore pre-adjusted such that the variances are on the same scale. Variance components estimated using the adjusted records were higher than using unadjusted records. Convergence of breeding value estimation is reached significantly faster when using adjusted data (± 4000 iterations) compared to unadjusted records (± 15 000 iterations). Although cow and bull rankings were not influenced much, significant changes in breeding values for individual animals and genetic trends of especially young animals, were found.

Keywords : Ayrshire, Guernsey, Holstein, Jersey, production traits, repeatability model, somatic cell score

#Corresponding author. E-mail : Bernice@arc.agric.za

7.2 INTRODUCTION

Considerable interest has been shown in the last few years to model individual test-day records for genetic evaluation of dairy cattle, instead of using the traditional accumulated 305-day yields in the evaluation of production traits (Swalve, 2000, Jensen, 2001 and Mrode *et al*., 2002). South Africa implemented test-day models for genetic evaluations of production traits, using a Fixed Regression Test-Day Model (FRTM), where the lactation curve is modeled as a fixed regression and the random components are specified as a traditional repeatability model, *i.e.* constant additive genetic and permanent environmental variances throughout the lactation. Test-day records of the first three lactations were included as repeated measures in the South African FRTM (Mostert *et al*., 2006a). A FRTM assumes equal variances at different days in milk, but data at the beginning and at the end of lactations have had higher variance than tests in the middle of lactation (Jensen, 2001). This is probably due to the onset and end of lactation processes being influenced by more factors than maintenance of production in the middle of lactation (López-Romero *et al*., 2003). For example, the interval between calving and first test of the lactation influences first test-day yield more than other test-day yields of the lactation, as can be expected in view of rapid changes in milk yield during early lactation (Pander *et al.*, 1992). Furthermore, first lactations have lower mean and variances compared to second and third lactations. Swalve (2000) and Jensen (2001) recommended that heterogeneity of variance should be accounted for in the application of test-day models.

In the test-day model implemented by Mostert *et al.* (2006a), it was found that older sires, especially those born during the 1980s, who performed well at their time, but whose dams' and other female relatives' test-day records were not stored on the INTERGIS (Integrated Registration and Genetic Information System) and who did not have active progeny in the TDM anymore, ranked too high (Mostert *et al.*, 2006b). The aim of this study was therefore to investigate adjusting for heterogeneous variances which exists at different days in milk and parity; to determine the effect of deviating from this assumption as required by repeatability models and to investigate the inclusion of a fixed calving year effect in the South African FRTM for genetic evaluation of production traits of Jersey cattle.

7.3 MATERIALS AND METHODS

A total of 3 192 159 milk, butterfat and protein test-day records of the first three parities for Jersey cows participating in the South African Milk Recording Scheme were downloaded from the INTERGIS. The following edits were applied to the data : deletion of records with unknown herds, unknown birth dates and calving dates, test days recorded before five days in milk or after 305 days in milk, records of crossbred cows and age restrictions within lactations to ensure reasonable calving ages in a specific lactation (17 – 40 months for lactation 1, 29 – 53 months for lactation 2 and 41 – 67 months for lactation 3). Protein yield was treated as missing for records where protein percentage was greater than 6% or less than 2%. The same was done with butterfat yield where butterfat percentage was higher than 9% or lower than 2%. Test-day milk yield was limited to a range of 1kg to 70kg. Lactations should fit the following requirements to be included (specifications from IRIS - the national dairy management system) :

1. First test of a lactation should be within the first 63 days in milk.
2. No interval longer than 100 days between tests of a lactation are allowed.
3. Only one interval between 60 and 100 days allowed per parity.

Lactations ending before 60 days in milk were also discarded. These specifications resulted in 13% of the data being discarded. A dataset of 2 768 524 test-day records was obtained. This dataset is referred to as the unselected dataset.

The following fixed effect model was then applied to the data, separately for each trait, in order to obtain least square solutions (LSS) of all fixed effects in the model :

$$y_{ijklmnp} = \mu + HTDLM_{im} + S_{km} + AC_{lm} + wilmink(S_{km}) + CI_{jm} + CY_{pm} + e_{ijklmnp}$$

Where

$y_{ijklmnpq}$ = p^{th} test-day milk, butterfat or protein yield of cow j in lactation m in herd x test-date x parity x number of milkings group i, of season k, age class l, calving interval class n and calving year p

μ = mean yield

$HTDLM_i$ = fixed effect of herd x test-date x parity x number of milkings group

S_{km} = fixed effect of calving season in lactation m

AC_{lm} = fixed effect of ageclass in lactation m

$wilmink(S_{km})$ = Wilmink curve (Wilmink, 1987) modeled on days in milk within season k and in lactation m (regression)

CI_{nm} = fixed effect of calving interval class in lactation m

$CYpm$ = fixed effect of calving year in lactation m

$e_{ijklmnp}$ = random residual error

Two seasons were defined : April – September and October – March, while the same age classes were allocated as in the derivation of standard lactation curves by Mostert *et al.* (2001). Calving interval classes were allocated using standard deviation units. This model assumed consecutive test day samples of a cow, within and across lactations, to be repeated observations of the same trait (Mostert *et al.*, 2006a).

Data was then pre-adjusted for heterogeneous variances as follows : Variances of residuals from the fixed effect model at each day in milk (DIM) were calculated separately for each parity, for all traits. E.g. variance of DIM i of lactation m is var_im. A weighted average of all var_im values was then calculated using SAS (1996) to obtain var_m, the average variance within lactation m. It was decided to use lactation 1 as the reference parity, as most test-day records originate from lactation 1 (1 246 080 lactation 1 records *versus* 1 016 606 for lactation 2 and 809 323 for lactation 3). The following scaling factor (s_im) was then implemented to pre-adjust all test-day records such that residual variance of all lactations and all days in milk were similar to the weighted average of lactation 1 (reference parity) (Dr Zenting Liu, 2005: Personal Communications, VIT Geneticist, zenting.liu@vit.de) :

$$\text{s\_im} = \sqrt{}\,(\text{var\_m} / \text{var\_im}) * \sqrt{}\,(\text{var\_1} / \text{var\_m})$$

$$= \sqrt{}\,(\text{var\_1} / \text{var\_im})$$

where var_1 is the weighted average of days in milk variances for first lactation (reference parity) and var_im is the variance of day i in milk in lactation m. After the estimation of scaling factors, the test-day records were adjusted as follows :

$$y^* = \text{LSSs} + \text{r*s\_im}$$

where y^* is the test-day yield adjusted for heterogeneous variances, LSSs are the least square solutions of all fixed effects in the model for the specific test-day and r is the residual variance. The adjusted yields (y^*) were then included in (co)variance component estimation and in the national genetic evaluation for estimation of breeding values, using the same model as above, but adding the animal additive genetic and permanent environmental effects as random effects.

For (co)variance estimation, a selected dataset was carefully constructed to ensure adequate genetic linkage amongst contemporary groups, as follows:

- Only records from cows where both parents were known.
- Only records where milk, butterfat and protein yields were measured.
- Cows must have a first parity.
- Contemporary groups with daughters of at least two sires.
- Contemporary groups with at least five records.
- Sires must be represented in at least three contemporary groups.
- Lactations must have at least nine test-days.

(Co)variance components were estimated with a multitrait analysis using VCE4 (Groeneveld & Garcia-Cortes, 1998). For this selected dataset, pedigrees were traced back for three generations.

PEST (Groeneveld & Kovac, 1990) was used to estimate breeding values, using the unselected dataset and a multitrait evaluation. The pedigrees were, however, traced back as far as possible and genetic groups were incorporated to ensure that base animals enter the evaluation on the appropriate genetic level.

Pearson correlations were estimated using SAS (1996) between adjusted test-day records and test-day records from the March 2005 national genetic evaluation (unadjusted) (Mostert *et al.*, 2006a), as well as between breeding values based on the adjusted records and breeding values from the March 2005 national genetic evaluation. Differences between breeding values (EBVs) from these two evaluations (adjusted breeding values – unadjusted breeding values), averaged per year of birth, were also plotted for proven sires (having at least 20 daughters in 10 herds), young sires (having at least one daughter and having been born since 1999) and measured cows. Cows were separated into cows having only first lactations, cows having up to the second lactation and cows having up to three lactations. Genetic trends from the adjusted as well as unadjusted evaluations, averaging EBVs per year of birth, for proven and young sires, as well as measured cows, were also calculated and compared. The base year adaptation was not done on these trends.

7.4 RESULTS AND DISCUSSION

Pre-adjustment of test-day records for heterogeneous variances are often done in genetic evaluations. Reents *et al.* (1998) describes an adjustment procedure to account for within herd heterogeneous variances, considering number of contemporary records, production levels, parity and stage of lactation, for official implementation in the previously used German Fixed Regression Test-Day Model. In the Canadian Random Regression Test-Day Model, data is pre-adjusted for heterogeneous herd-test-date-parity variances on a trait to trait basis (Schaeffer *et al*., 2000).

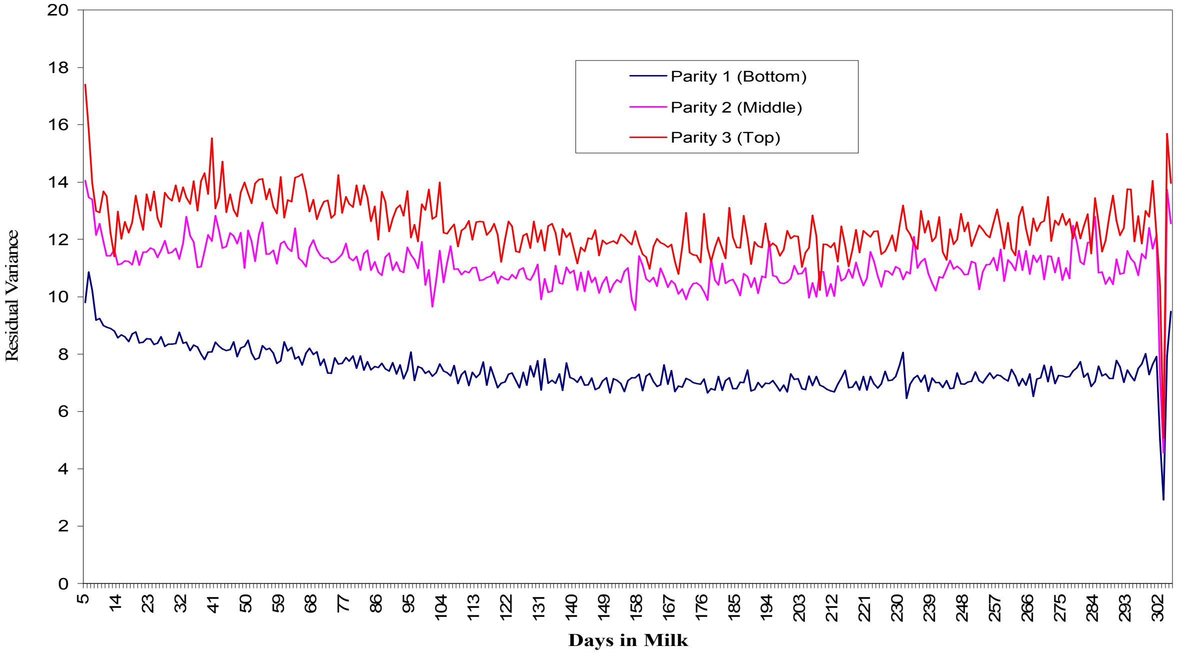


Figure 7.1 Residual variance per day in milk for lactations 1, 2 and 3 of the Jersey breed.

Figure 7.1 indicates residual variances across the lactation. It is clear that residual variances are higher in the beginning of all lactations compared to tests in the middle of the lactation, with erratic behavior, due to only few tests available, at the end of the lactations. Residual variances of first lactations are dramatically lower during all stages of the lactation compared to second and third lactations, while that of second and third lactations are more comparable. Third lactation test-day records show the highest variance throughout the lactation. Adjusting for heterogeneous variance due to days in milk and parity will therefore render an improvement on the March 2005 national genetic evaluation.

In Table 7.1 the data structure and statistics of the datasets for (co)variance component estimation and for prediction of breeding values, are presented. From this table it can be seen that 3.75% of the data was selected for (co)variance component estimation. As a higher proportion of first lactation records (77%) relative to second and third lactation records was included in the selected dataset compared to the unselected dataset (42% first lactation records), averages in the selected dataset were slightly lower for all traits compared to the unselected dataset.

Table 7.1 Data structure and statistics (average ± SD) of the Jersey datasets used for (co)variance component estimation (selected dataset) and prediction of breeding values (unselected dataset).

	Unselected Dataset	**Selected Dataset**
Test-Day Records	2 768 524	103 889
Lactation Records	361 352	11 343
Cows	190 372	8 717
Sires	4 504	949
Dams	105 406	7 681
Contemporary Groups	181 411	7 423
Pedigree	334 300	23 651
Milk Yield (kg/day)	15.65 ± 5.28	14.42 ± 4.50
Butterfat Yield (kg/day)	0.71 ± 0.26	0.65 ± 0.21
Protein Yield (kg/day)	0.58 ± 0.19	0.54 ± 0.16

Pearson correlations between unadjusted (March 2005 national genetic evaluation) and adjusted test-day records, were above 99% for all traits, with butterfat yield mostly affected.

In Table 7.2 variance component ratios estimated based on adjusted records from the selected dataset, as well as variance components used in the March 2005 national genetic evaluation (unadjusted records), are listed.

Table 7.2 Variance component ratios estimated using the selected dataset, adjusted for heterogeneous variances and unadjusted data, as used in the March 2005 national genetic evaluation, for the Jersey breed.

	ADJUSTED DATA			**UNADJUSTED DATA**		
			DIRECT			
	Milk	**Butterfat**	**Protein**	**Milk**	**Butterfat**	**Protein**
Milk	0.242 ± 0.017	0.774 ± 0.019	0.912 ± 0.008	0.175 ± 0.013	0.779 ± 0.023	0.898 ± 0.010
Butterfat		0.200 ± 0.013	0.884 ± 0.011		0.103 ± 0.009	0.882 ± 0.014
Protein			0.229 ± 0.017			0.155 ± 0.012
			PERMANENT ENVIRONMENT			
Milk	0.374 ± 0.016	0.923 ± 0.007	0.965 ± 0.003	0.304 ± 0.012	0.886 ± 0.008	0.952 ± 0.003
Butterfat		0.262 ± 0.012	0.943 ± 0.005		0.193 ± 0.008	0.932 ± 0.006
Protein			0.329 ± 0.015			0.259 ± 0.011
			RESIDUAL			
Milk	0.384 ± 0.003	0.752 ± 0.001	0.912 ± 0.000	0.521 ± 0.004	0.723 ± 0.001	0.918 ± 0.000
Butterfat		0.538 ± 0.004	0.741 ± 0.001		0.703 ± 0.004	0.708 ± 0.001
Protein			0.442 ± 0.004			0.586 ± 0.004

Variance component ratios estimated using the adjusted records were higher for the direct and permanent environmental effects, at the expense of residual ratios, in comparison with estimates obtained from unadjusted data. This makes sense as part of the residual variance has already been taken care of when adjusting for heterogeneous variances. However, these (co)variance component estimations were not done on the same dataset. Different criteria were used in the selection of the two datasets used in these estimations. For example in the adjusted dataset, one of the criteria was that only lactations with at least nine test-day records were included in the selection, while in the unadjusted data set, lactations consisting of six and more test-day records, were included (Mostert *et al*., 2006a).

Variance component estimation on test-day records is influenced by the stage of the lactation which is represented by the test-day records included in the evaluation. According to Meyer *et al.* (1989) heritabilities were generally highest for test-day yields in the second trimester of lactation. Pander *et al.* (1992) reported that heritability estimates for milk yield, butterfat and protein concentrations were highest in mid-lactation with a similar pattern for butterfat and protein yields, except that estimates in late lactation for these yields did not fall. Jakobsen (2000) showed that for milk and protein yield, there was a tendency towards higher heritability estimates in mid-lactation, while heritabilities for butterfat yield were more constant throughout the lactation. Druet (2003) also found genetic variance to be highest in mid-lactation and lower at the beginning and end of lactation. This can be attributed to intervals being too wide between test-days in the extremes of lactation to define the traits in those parts of the lactation and because information is scarce in those periods (López-Romero *et al.*, 2003). Pre-adjusting test-day records for heterogeneous residual variances removed the decline in residual variance that occur throughout the lactation, as well as the slight increase at the end of lactation, as described by Meyer *et al.* (1989), Pander *et al.* (1992), Swalve (1995), Rekaya *et al.* (1999) and Pool *et al.* (2000), rendering higher (co)variance component ratios in comparison with those based on the unadjusted records which included a higher percentage lactations over a shorter stage of the total lactation.

Convergence of breeding value estimation (defined as standardized maximum change of the solutions from one round to the next with a stopping criterion of 0.001) was reached much faster when using adjusted data (± 4000 iterations) compared to unadjusted records (± 15 000 iterations for the March 2005 evaluation). Correlations between EBVs on adjusted and unadjusted records were higher than 96% for measured cows and young sires, and higher than 98% for proven sires.

Figure 7.2 Differences in milk yield EBVs (adjusted-unadjusted), averaged per year of birth, for proven and young Jersey sires.

Figure 7.2 shows the differences in milk yield EBVs (adjusted evaluation EBVs – unadjusted evaluation EBVs) averaged per year of birth, for young and proven sires. For young sires EBVs were overestimated by the March 2005 evaluation, with the youngest sires (having a larger proportion of their daughters early in their lactation), being affected mostly. EBVs of proven sires, born from 1983 to 1993, were also overestimated by the March 2005 evaluation, with sires born in 1985 especially affected. The lower EBVs estimated by the adjusted evaluation are probably due to the inclusion of the calving year effect in the model, which proved to be significant for all traits (P<0.05). Averages of sires born before 1979 were based on only a few sires per year.

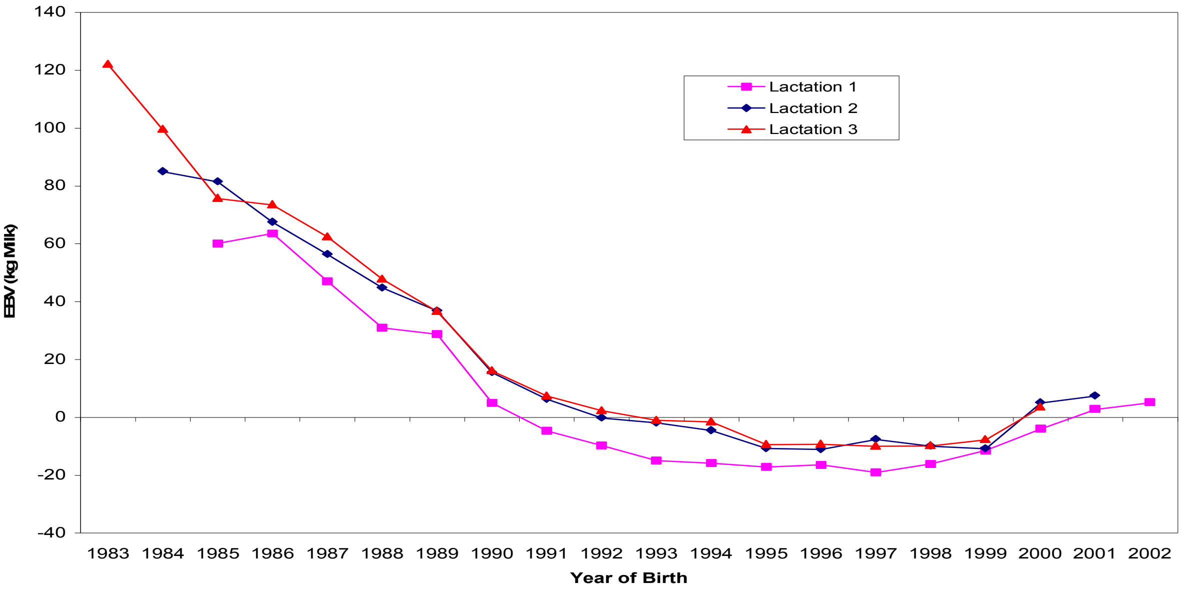


Figure 7.3 Differences in milk yield EBVs (adjusted-unadjusted), averaged per year of birth, (adjusted - unadjusted), for Jersey cows.

In Figure 7.3 these differences are indicated for the cows. EBVs were underestimated by the March 2005 evaluation for cows born between 1983 and 1990 (first lactation cows) and 1992 (second and third lactation cows), with second and third lactation cows mostly affected. Thereafter EBVs were overestimated, with first lactation cows being mostly affected, followed again by a slight underestimation for the young cows.

These changes were, however, indeed expected. In particular young bulls with a larger proportion of their daughters early in their lactation were affected most, because the test-day data from the early lactation stages, which have higher phenotypic variances than in the middle of lactation, were regressed more towards the reference base as a result of the adjustment for heterogeneous variances due to days in milk and parity. For the same reason, young cows with few test-day records should be affected more than cows with complete lactations.

Table 7.3 Summary of statistics for individual differences (adjusted milk yield EBVs – unadjusted milk yield EBVs) of young sires, proven sires and first lactation cows, cows having up to second lactations and cows having up to third lactations.

	Number	**Min. Difference[1] (kg)**	**Max. Difference [2] (kg)**	**Avg. Difference (kg)**	**SD**
Young Sires	365	-888	530	-11.2	134.9
Proven Sires	418	-180	228	-22.8	63.9
1st Lactation Cows	46 649	-1 157	1 157	-1.5	98.3
2nd Lactation Cows	43 642	-719	861	8.1	71.8
3rd Lactation Cows	95 649	-886	725	14.1	64.8

[1]Overestimated [2]Underestimated

Although bull and cow rankings as well as differences between the two analyses, averaged per year of birth, were not dramatic, significant changes for individual animals were found between the two models. Table 7.3 indicates that the average EBVs of proven and young sires and first lactation cows were lower after adjusting for heterogeneous variances and fitting a calving year to the model, therefore these EBVs were previously on average, overestimated. Second and third lactation cows' averages, were higher after these changes to the model. These EBVs were therefore previously, on average, underestimated. Individuals that differed most, were first lactation cows, with proven sires being the least affected.

For example, individual sires were found with milk EBVs being 180kg lower (previously overestimated) and 228kg higher (previously underestimated) compared to EBVs from the March 2005 evaluation. EBVs of young sires showed by far the most variation after the changes in the model. Reents *et al.* (1998) also reported that although overall cow and bull rankings were not influenced much by method of standardization for heterogeneous herd x test-date variances, significant effects for individual animals could be found.

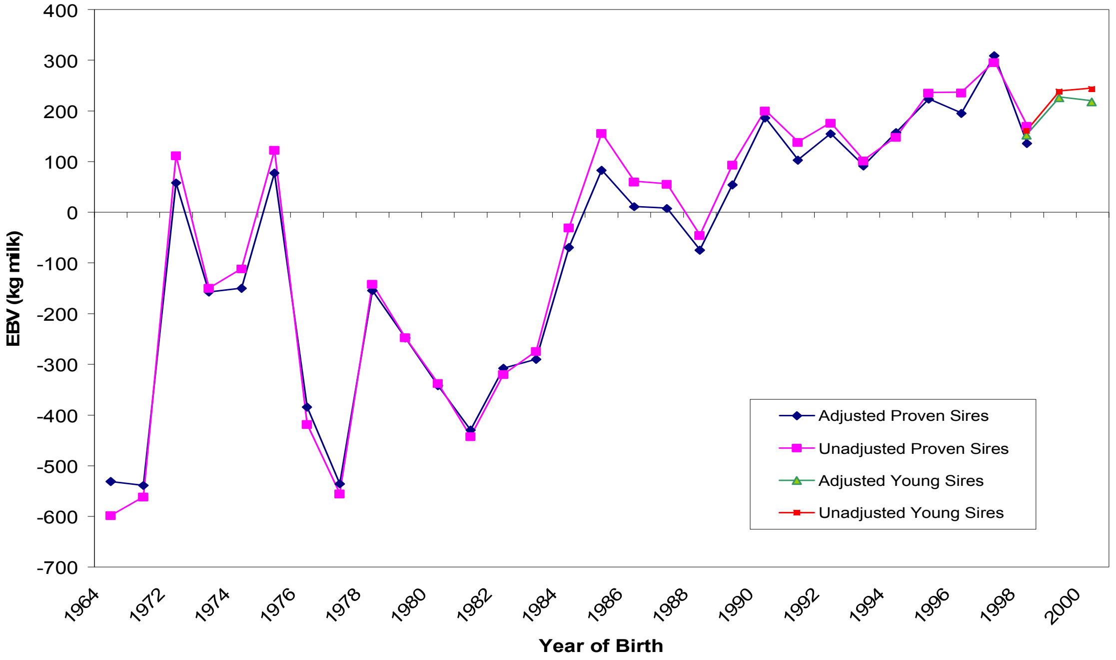


Figure 7.4 Effect of adjustment for heterogeneous variances and inclusion of calving year in the model, on the milk yield genetic trend of proven and young sires.

Figures 7.4 and 7.5 indicate the effect of adjustment for heterogeneous variances and the inclusion of a fixed calving year effect in the model on the genetic trend of milk yield for sires and cows, respectively. The genetic trend for proven sires is generally lower for the adjusted evaluation, especially for sires born from 1983 to 1993 and again in 1996. For the young sires the adjusted evaluation also yielded a lower genetic trend, with the youngest sires showing the largest decline in trend. For the cows, adjustment and adding a fixed calving year effect to the model only influenced the younger cows, with the adjusted evaluation yielding a significantly lower trend for cows born after 2001. Similar tendencies were observed for butterfat and protein yields, resulting therefore in a lower percentage of young animals being in the top bull and cow rankings compared to the unadjusted evaluation.

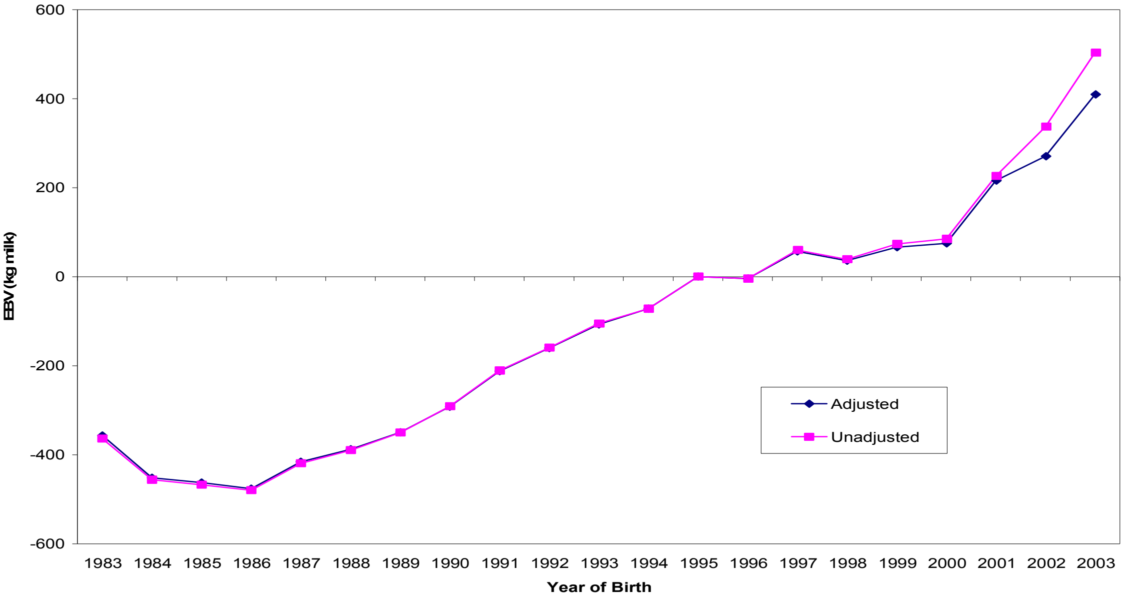


Figure 7.5 Effect of adjustment for heterogeneous variances and inclusion of calving year in the model, on the milk yield genetic trend of cows.

Reents *et al.* (1998) also observed that highest genetic trends were found for models without any adjustment and smallest with models with a strict adjustment for heterogeneous variances.

7.5 CONCLUSIONS

Pre-adjusting for heterogeneous variances due to DIM and parity has definite advantages with regards to convergence time for national evaluations. Although cow and bull rankings were not influenced much, significant effects for individual animals and genetic trends of especially young animals were found. Including a fixed calving year effect and adjusting for heterogeneous variances due to days in milk and parity result therefore in more accurate estimation of breeding values, especially for young animals. It is recommended that these changes should be implemented in the national genetic evaluations of the other dairy breeds, thereby preventing the over- and underestimation of individual EBVs.

7.6 REFERENCES

Druet,T., Jaffrézic, Boichard, D. & Ducrocq, V., 2003. Modeling lactation curves and estimation of genetic parameters for first lactation test-day records of French Holstein cows. J. Dairy Sci. 86, 2480-2490.

Groeneveld, E. & Garcia-Cortes, A., 1998. VCe4.0, a (co)variance component package for frequentists and Bayesians. Proc. 6th WCGALP. 27, 455-456.

Groeneveld, E. & Kovac, M., 1990. A generalized computing procedure for setting up and solving mixed linear models. J. Dairy Sci. 73, 513-531.

Jakobsen, J.H., 2000. Genetic correlations between the shape of the lactation curve and disease resistance in dairy cattle. Ph.D. Thesis. Dept. of Animal Breed. Genet. Danish Institute of Agricultural Sciences, Research Centre, Foulum.

Jensen, J., 2001. Genetic Evaluation of Dairy Cattle using Test-Day Models. J. Dairy Sci. 84, 2803-2812.

López-Romero, P. & Carabano, M.J., 2003. Comparing alternative random regression models to analyse first lactation daily milk yield data in Holstein-Friesian cattle. Livest. Prod. Sci. 82, 81-96.

Meyer, K., Graser, H.-U. & Hammond, K., 1989. Estimates of genetic parameters for first lactation test day production of Australian Black and White cows. Livest. Prod. Sci. 21, 177-199.

Mostert, B.E., Theron, H.E. & Kanfer, F.H.J., 2001. The effect of calving season and age at calving on production traits of South African dairy cattle. S. Afr. J. Anim. Sci. 31, 205-214.

Mostert, B.E., Theron, H.E., Kanfer, F.H.J. & Van Marle-Köster, E., 2006a. Fixed regression test-day models for South African dairy cattle for participation in international evaluations. S. Afr. J. Anim Sci, submitted.

Mostert, B.E., Theron, H.E., Kanfer, F.H.J. & Van Marle-Köster, E., 2006b.Comparison of breeding values and genetic trends for production traits estimated by a Lactation Model and a Fixed Regression Test-Day Model. S. Afr. J. Anim Sci, 36, 71-78.

Mrode, R.A., Swanson, G.J.T. & Lindberg, C.M., 2002. Efficiency of part lactation test day records for genetic evaluations using fixed and random regression models. Anim. Sci. 74, 189-197.

Pander, B.L., Hill, W.G. & Thompson, R., 1992. Genetic parameters of test day records of British Holstein-Friesian heifers. Anim. Prod. 55, 11-21.

Pool, M.H., Janss, L.L.G. & Meuwissen, T.H.E., 2000. Genetic parameters of Legendre polynomials for first parity lactation curves. J.Dairy Sci. 83, 2640-2649.

Reents, R., Dopp, L., Schmutz, M. & Reinhardt, F., 1998. Impact of application of a test day model to dairy production traits on genetic evaluations of cows. Proc. 1998 INTERBULL Meeting, Rotorua, New Zealand, January 18-19, 49-54.

Rekaya, R., Carabano, M.J. & Toro, M.A., 1999. Use of test day yields for the genetic evaluation of production traits in Holstein-Friesian cattle. Livest. Proc. Sci. 57, 203-217.

SAS, 1996. SAS user's guide : Statistics, Release 6.12. SAS Institute Inc., Cary, North Carolina, USA.

Schaeffer, L.R., Jamrozik, J., Kistemaker, G.J. & Van Doormaal, B.J., 2000. Experience with a Test-Day Model. J. Dairy Sci. 83, 1135-1144.

Swalve, H.H., 1995. The effect of test day models on the estimation of genetic parameters and breeding values for dairy yield traits. J. Dairy Sci. 78, 929-938.

Swalve, H.H., 2000. Theoretical basis and computational methods for different test-day genetic evaluation methods. J. Dairy Sci. 83, 1115-1124.

Wilmink, J.B.M., 1987. Adjustment of test-day milk, fat and protein yield for age season and stage of lactation. Livest. Prod. Sci. 16, 335-348.