

Two years of experience with the French genetic evaluation of dairy bulls on production-adjusted longevity of their daughters

Vincent Ducrocq

Station de Génétique Quantitative et Appliquée,
Institut National de la Recherche Agronomique,
78352 Jouy en Josas, France

Abstract

Four runs of the routine genetic evaluation of French dairy bulls on production-adjusted length of productive life of their daughters have been released since June 1997. This paper describes the current model of evaluation and how and why it has changed over the past two years. The computation of reliabilities is explained in details and the use of an approximation of the heritability of length of productive life on the observed scale is justified. Main results for fixed effects and relationships of sire effects with other traits are presented. A recurrent problem is the existence of strong upward biases in the evaluation of top proven sires with young batches of second daughters. This phenomenon, combined with the absence of reliable evaluations for young sires, explain a certain lack of concern for the new evaluation on behalf of breeders. Ways of improvements are indicated.

1. Introduction

In France, the first genetic evaluation of dairy bulls on production-adjusted length of productive life of their daughters was released in June 1997. The routine evaluation is done twice a year, with release in June and November. With respect to the first implementation, a few changes in the way censored records are defined and in the model were introduced. The motivation of these changes, the problems encountered and some interesting and/or unexpected results are described in this paper.

2. The current implementation

2.1 Data sets and censoring definition

The last available evaluation was run in October 1998 for a release in November. The datasets analysed were directly extracted from the datasets used for the routine production traits evaluation of late September 1998. They included all cows milk recorded between December 1, 1984 and May 1, 1998, and whose sire had at least 20 daughters (5 for small breeds). The latter requirement may seem very strict. Its origin is directly related to computational considerations: as it will be seen, convergence is slow and computing time is a limiting factor. Furthermore, sires with fewer than 20 daughters in the production traits evaluation are either natural service bulls, very old bulls or young AI bulls with

not enough information to have an official proof for production and functional traits.

For cows with first calving before December 1984, only the part corresponding to lactations started after that date was used (« left-truncated » records; truncation date = first calving date after December 1, 1984).

In France, until the end of 1998, exit dates from the herds were not systematically recorded nor reliable. Censoring codes (i.e., indicator variable of culling) have to be attributed indirectly, from the existence or the absence of a test day record in the file, after given dates. More precisely, three categories of records are censored :

- Records of cows with last test date (τ_{last}) after $\tau_{\text{max}} = \text{May 1, 1998}$ are considered as censored at τ_{max} . The choice of this censoring time is critical: in the very first run, records from cows still alive at $\tau_{\text{last}} > \tau_{\text{max}}$ were censored at τ_{last} . As a consequence, such cows were never at risk between τ_{max} and τ_{last} (they would never get an uncensored record) although for some cows, the test date τ_{last} was the last before culling (unknown at the time of analysis). In the evaluation, these extra ($\tau_{\text{last}} - \tau_{\text{max}}$) days were « credited » to the sires. For a proven sire with a massive arrival of second crop daughters, this extra time multiplied by a very large number of cows (e.g., several thousands) resulted in big favourable biases (up to 1 or 2 genetic standard deviations).

- Only the first part of records of cows changing herds during their productive life is included (and censored). The second part is discarded, as the attitude of the farmer towards such cows is very likely to be atypical.

- Records in herds such that the herd size is decreasing by more than 50% in one year are censored at the beginning of the relevant year-season. Such herds are either going out of test or suffering a massive disease outbreak, which make them no longer representative of the population. When such records are not properly treated as censored, the herd-year-season variance is strongly and unduly inflated.

2.2 Current model

The initial model used in June 1997 is described in Ducrocq and Sölkner (1998a). The current one differs slightly. It is a Weibull frailty model with time-dependent covariates, defined on three different time scales: the function $h(t)$ represents the hazard of a cow, at calendar time τ , t days after her first calving, t' days after her last calving is :

$$h(t) = \rho t^{\rho-1} * \exp\{y_i(\tau) + hy_r(\tau) + I_j(t, t') + a_1 + v_q(\tau) + m_k(t) + ym_{ik}(t) + f_n(t) + p_o(t) + s_u + 0.5s_{gs}\} \quad [1]$$

where:

- ρ is one of the parameters of the Weibull baseline hazard function (the other one is implicitly included in the exponential part and plays the role of a grand mean). Preliminary analyses showed that a fixed value of $\rho=2$ was convenient;

- $y_i(\tau)$ is the time-dependent year-season effect at calendar date τ , with changes on March 1 and December 1 each year. The choice of these dates is driven by the belief that culling policies change at the end of the quota period (April 1st to March 31st);

- $hy_r(\tau)$ is the time-dependent random herd-year-season effect at τ with changes on March 1 and December 1 each year.

Its distribution is assumed to be log-gamma(γ, γ). Preliminary results showed that a constant value of $\gamma = 4$ can be used (fixing ρ and γ is computationally advantageous). In the evaluation process, herd-year-season effects are integrated out and therefore are not explicitly computed. Note that the simultaneous inclusion of $y_i(\tau)$ and $h_r(\tau)$ is a way to specify that the mean of the random herd-year-season effect can vary with time;

- $I_j(t, t')$ is the time-dependent combined effect of lactation number, t days after first calving (lactations 1, 2, 3, 4, 5, 6+) and of stage of lactation, t' days after the current calving (with changes at $t'=30, 60, 150, 240$ and date when dried). A cow is considered as dried between her last test day of the lactation and her next calving date. Since she must have a next

calving to be defined as dried, she is not at risk of being culled during the dry period. This definition of a dry period class annihilates the otherwise favourable effect of long dry periods on length of productive life. Note that an undesirable effect of long lactation lengths remains ;

- a_1 is the time-independent age at first calving with 23 levels (<21 mo, >41 mo and each month in between) ;

- $v_q(\tau)$ is the time-dependent combined effect of 4 herd size classes and (up to) 5 variation in herd size classes with changes on March 1 and December 1 each year;

- $m_k(t), f_n(t), p_o(t)$ are time-dependent classes of within herd-year deviations for milk yield (10 classes of equal size), fat content (5) and protein content (5) with changes at each new calving date ;

- $ym_{ik}(t)$ is the interaction between herd-year deviation for milk yield and year-season ;

- s_u and s_{gs} are the random transmitting abilities of the sire and of the maternal grand-sire of the cow. When the maternal grand-sire is not known, the term $0.5s_{gs}$ is simply ignored. These effects are grouped into a vector \mathbf{s} which follows a multivariate normal distribution with variance-covariance matrix $\mathbf{A} \sigma_s^2$. Genetic parameters are given in table 1.

Compared to the initial model (Ducrocq and Sölkner, 1998a), 3 important changes were included : the model is now a sire-maternal grand-sire model instead of a sire model. An age at first calving was added as well as the interaction between milk production and year-season : global changes in voluntary culling policy for production traits over time are now accounted for.

The Normande, Montbeliarde and Holstein breeds are analysed separately. For the 12 other dairy breeds of much smaller population size, a joint analysis is performed using the sire variance of the Normande breed and with a modified model: a breed effect is included; age at first calving and the interaction between milk production and year-season are ignored; effects of year-season, of lactation number x stage of lactation, of herd-year deviations of production traits and of variation in herd size are defined within breed. If a sire is used in more than one breed, within breed sire effects are computed. Data sets characteristics are summarised in table 1.

2.3 Computational considerations

The evaluation is performed using the « Survival Kit-V3.0 » (Ducrocq and Sölkner, 1998b) on an IBM Risc 595 AIX 4.21 computer with 2 Gigabytes of memory. The inclusion of many time-dependent covariates in the model results in an average of 19.4

elementary records per cow. Hence, for the Holstein breed, 134 million elementary records are created. This is equivalent to an unformatted input file of size 15 Gigabytes. The necessity to cope with such huge files was the primary motivation to write the modified programs *preparec.f* and *weibullc.f* of the Fortran programs *prepare.f* and *weibull.f* in version 3.0 of the Survival Kit. These programs use public domain C subroutines for compressing and decompressing data during I/O operations. The size of the compressed recoded file for the Holstein breed was 766 Megabytes, i.e. about 20 times smaller than the uncompressed file. In practice, compression (once) and decompression (at each iteration) multiply by 3 the overall computing time, with respect to an uncompressed situation.

In contrast with regular genetic evaluations, I/O operations are never slowing down CPU. The limiting factors are the compression / decompression steps and the calculation of one exponential function for *each* elementary record at *each* iteration. For the three major breeds, starting values equal to solutions of the previous evaluation are used, including for sire effects (with the use of a particular preparation program). 200 to 300 iterations are run. CPU times are given in table 1.

Monitoring of convergence is difficult. However, it was repeatedly verified that sire solutions were not varying by more than 0.1 genetic standard deviation over the last 50 iterations.

Table 1. Datasets characteristics, genetic parameters and computing considerations

	Montbéliarde	Normande	Holstein	Other breeds
Records (total)	925537	1057326	6897881	155074
Censored (%)	31.4	26.6	28.0	29.1
Left truncated (%)	12.4	11.3	11.8	14.3
Sires	3516	4350	18688	4686
Sire variance	0.047	0.039	0.053	0.039
h ² (% , original scale)	19.3	16.1	21.7	16.1
Size of recoded file (Mbytes)	123	126	766	22
Data preparation	17'	18'	1h50'	
Recoding	35'	37'	4h08'	
Evaluation	8h55' ⁽¹⁾	9h30' ⁽¹⁾	79h30' ⁽²⁾	1h58' ⁽³⁾
Final convergence criterion ⁽⁴⁾	1.10 ⁻⁸	7.10 ⁻⁹	4.10 ⁻⁸	4.10 ⁻⁸

^{1,2,3} starting from previous solutions (except for ⁽³⁾) and after 200, 250 and 300 iterations respectively.

⁴ as defined in « the Survival Kit »

3. Publication of estimated breeding values and reliabilities

Estimated sire effects are standardised into genetic standard deviation units, with a change in sign:

$$\hat{g}_u = -\hat{s}_u / \sigma_s$$

so positive values are favourable (negative values of \hat{s}_u mean reduced culling). This scale is unit free and the same as for all other functional traits EBVs in France.

Together with the EBVs, it is important to supply breeders with a measure of accuracy of the evaluation. The usual definition of the reliability is based on the equation :

$$R_{true} = 1 - \frac{\text{asymptotic prediction error variance}}{\text{Var}(s)} \quad [2]$$

where the "exact" asymptotic prediction error variance is obtained from the diagonal term of the inverse of the information matrix (= - the Hessian of the log-likelihood function). Unfortunately, in large applications, this matrix is far too large to be calculated and inverted. Traditionally, an

approximate reliability $R_{appr.}$ is derived from selection index theory. In particular, if pedigree information is ignored and a sire model is used, the approximate reliability $R_{prog.}$ of the genetic evaluation of a sire with n progeny is:

$$R_{prog.} = \frac{n}{n + \frac{4-h^2}{h^2}} \quad [3]$$

With a survival model, with censored records and very skewed distributions, the question is: what value of h^2 should be use? (and subsequently, what value of n?). In the case of a Weibull mixed model such as [1], it is possible to define h^2 on the logarithmic scale, i.e., for log T as (Ducrocq and Casella, 1996):

$$h_{log}^2 = \frac{4\text{Var}(s_u)}{\text{Var}(\log T)} = \frac{4\sigma_s^2}{\sigma_s^2 + \psi^{(1)}(\gamma) + \frac{\pi^2}{6}} \quad [4]$$

where ω follows an extreme value distribution and $\Psi^{(1)}$ is the trigamma function (Kalbfleisch and Prentice, 1980). However, this heritability is difficult to interpret and does not seem to be related with the

reliability of evaluation. To compute an approximate reliability R_{appr} , an approximation h_0^2 of the heritability on the *observed* scale can be derived, using a Taylor series expansion around the mean of $\log T$. This derivation is described in details in the appendix.

Applied to the parameters of the French evaluation, this transformation leads to a h_0^2 about twice as large as h_{\log}^2 . This seems an obscure way to get larger heritabilities. However, we have repeatedly found that it leads to an excellent approximation of R_{true} . As an example, given the moderate size of the Normande breed dataset, it is possible to run the

evaluation and get the true asymptotic prediction error variance (and therefore R_{true}) in this breed. R_{prog} was calculated using formula [3] with various values of heritability and with n equal to either the total number of daughters or the number of uncensored daughters of each sire. R_{appr} was then obtained incorporating information from the sire's own sire (with reliability R_{gsire}) with the formula:

$$R_{appr} = \frac{\frac{1}{4}R_{gsire} + R_{prog} - 2 * \frac{1}{4}R_{gsire} * R_{prog}}{1 - \frac{1}{4}R_{gsire} * R_{prog}} \quad [5]$$

Figure 1: Comparison of true and approximate reliabilities (x 100) of Normande bulls when the approximate formula is $R=n/n+k$ with $k=(4-h^2)/h^2$ for different values of h^2 and n = all daughters or uncensored daughters (« true » values for h^2 in the Normande breed evaluation: 0.08 on the log-scale, 0.161 on the observed scale)

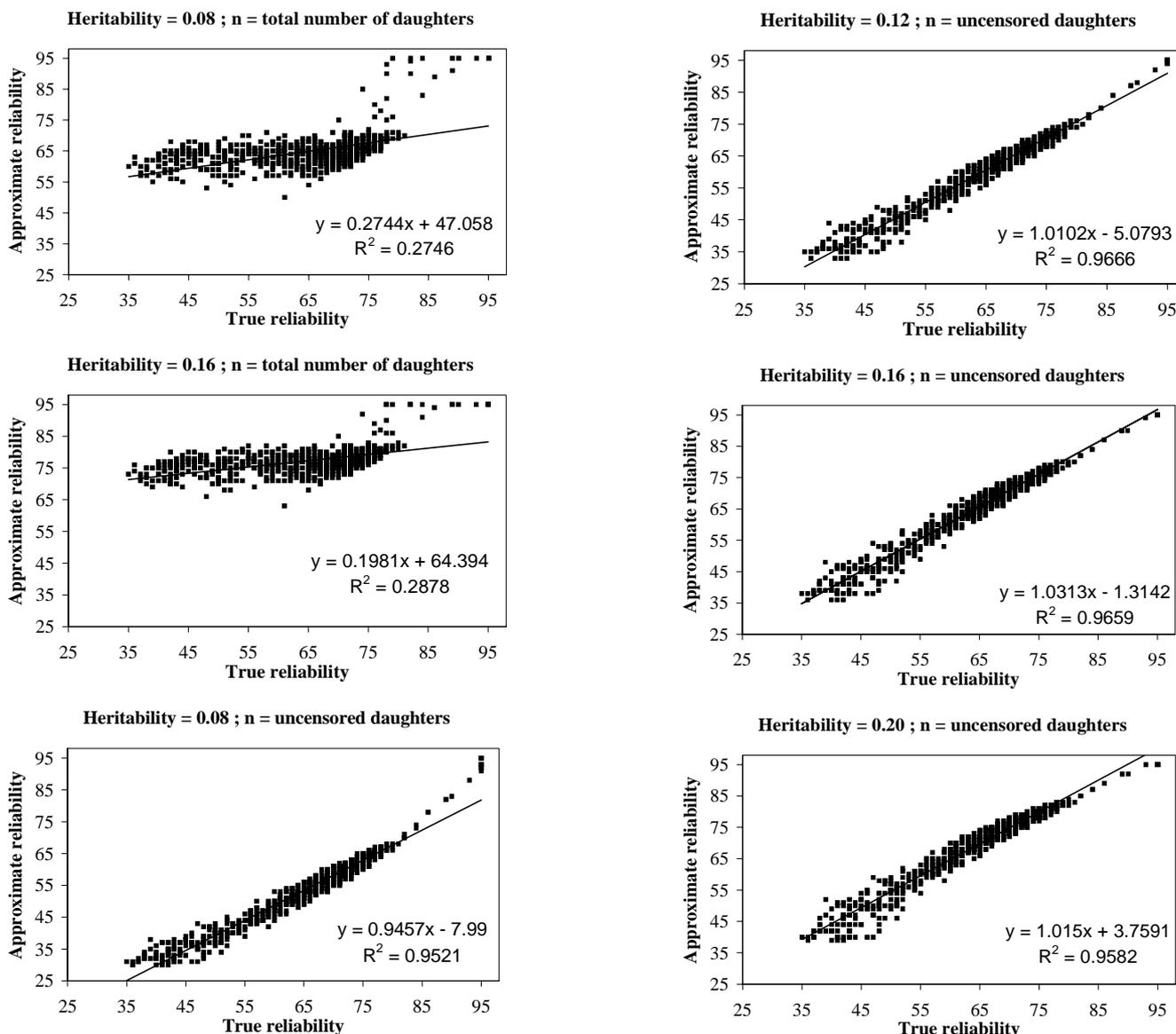


Figure 1 present 6 plots of $R_{appr.}$ vs R_{true} : clearly, $h^2 = h_0^2$ and $n =$ number of *uncensored* progeny give the best fit. For $R_{true} > 0.30$, $(R_{appr.} - \hat{R})$ ranges from -0.006 to $+0.015$, where \hat{R} refers to the value on the regression line of $R_{appr.}$ on R_{true} .

It must be strongly emphasised that this excellent “behaviour” of the approximation of h^2 on the observed scale may not be always found. Indeed, we have encountered two other situations (not related to the French evaluation) where h_0^2 did not make sense

at all ($h_0^2 > 1.0$); both situations were characterised by $\rho < 1$). This suggests that a thorough check of the validity of the propose approximation should be systematically performed.

In practice, only breeding values of bulls with $R_{appr.} > 0.50$ are published in France. When $R_{gsire} = 0.80$, this lower limit is obtained when a bull has 18, 15 or 14 *uncensored* daughters, for the Normande, Montbéliarde and Holstein breeds respectively. To put these numbers in perspective, note that this same accuracy can be obtained for a linear trait with $h^2 = 0.05$ with 59 to 62 *total* daughters. For a reliability $R_{appr.} = 0.70$, the corresponding figures are 50, 42 and 38 *uncensored* daughters, or 165 to 166 *total* daughters. This illustrates an important feature of the evaluation based on survival analysis: sires whose daughters are dying faster are more accurately evaluated than the others. It also underlines that h_0^2 is the heritability of length of productive life in total absence of censoring (i.e., when number of uncensored daughters = total number of daughters).

4. Results

4.1 Fixed effects

Obviously, it is not possible here to summarise results for all fixed effects. Only the most characteristic features for the three major breeds will be indicated:

- A systematic difference between relative risks in winter and summer is observed: just before the end of the quota period (winter), cows are at a 20-25% larger probability of being culled;
- Great care must be taken when one tries to interpret estimates of stage of lactation and lactation effects, as they are strongly associated with the time scale: one cannot compare a 1st and a 5th lactation cow at time, say, $t = 100$ days after first calving ! Indeed, it is more appropriate to look at an estimated hazard function:

$$\hat{h}(t) = \hat{\rho} t^{\hat{\rho}-1} * \exp\{ \hat{I}_j(t, t') \} \quad [6]$$

where $\hat{I}_j(t, t')$ includes the intercept (grand mean).

This calculation requires the definition of the record of a “reference cow”, in the sense that time of changes t and t' of the time-dependent covariate $I_j(t, t')$ must be known. Figure 2 represents the hazard of a Normande “reference cow” with constant 305 days lactations and 365 days calving intervals: within lactation, culling risk increases gradually and then sharply at the end. The hazard is very similar for the first two lactations, then consistently increases with time.

Figure 2: Hazard function of a reference cow with constant lactation length and calving interval, in the Normande breed

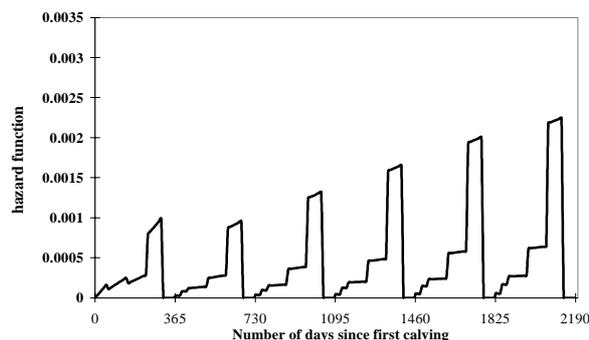
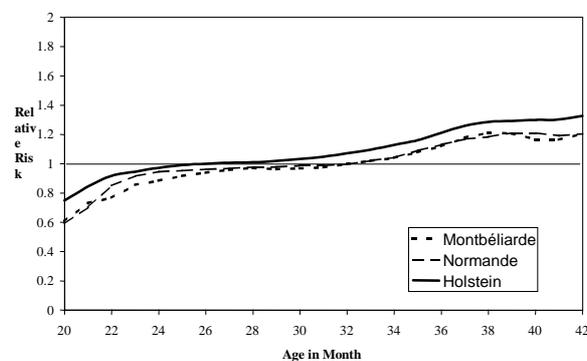


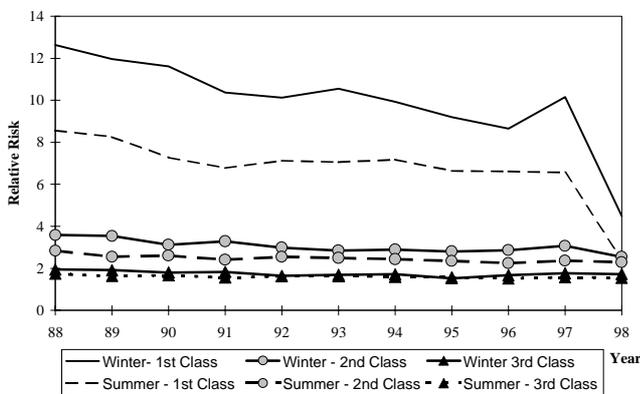
Figure 3: Estimates of age at first calving (reference for which relative risk = 1: 26 mo for the Holstein breed, 32 mo for the Normande and Montbéliarde breeds)



- Effects of age at first calving are almost identical for the three breeds (Figure 3). Very young first calvers (< 24 mo) are at a lower risk, mainly because they are less penalised for their lower milk production. Late calvers (> 34 mo) are at a higher risk. Between these two limits, age at calving has no effect.
- Cows in shrinking herds (by more than 15%) are at 60 to 80% higher risk of being culled than cows in stable herds. In expanding herds (by more than 15%), relative risk is reduced by 20 to 25%. Herd size has no real impact.
- Effect of within herd class of milk production and its interaction with year season are illustrated in

Figure 4: the risk of being culled is increased up to 13 times for cows in the bottom 10% (class 1) of the herd, compared to an average cow (class 5). But this higher risk is much more pronounced (+30 to +40%) in winter (before the end of the quota period) than in summer. It was also much more pronounced 10 years ago: it seems that voluntary culling for milk yield has declined over the years. The seasonal differences in voluntary culling (on top of the seasonal difference accounted by the year-season effect) and the time trend still exist but at a reduced level for the next 10% of the herd (class 2). For classes 3 and more, differences are no longer significant.

Figure 4: Estimates of the interaction of within-herd class of milk production and year-season, for the bottom 3 classes, in the Normande breed.



- Cows in the lowest class for protein percent are at a 70 to 100% higher risk of being culled than average cows. Best cows for protein content see their risk of being culled reduced by 18 to 34%. Fat content has very limited effect in the Holstein breed, but worst cows for fat content are at a 30 to 40% higher risk in the Montbéliarde and Normande breeds.

4.2 Sire effects

Computing survival curves for reference cows (see above) of sires with different breeding values allows a better visualisation of genetic differences: daughters of sires with a breeding value of 1 (genetic standard deviation) live on average about 160 days more (1/2 lactation). 5% (respectively 7%) more of them are still alive after 1 (resp. 2) lactations. Table 2 shows the correlation between the EBVs for production-adjusted length of productive life and other traits. These are underestimates of true genetic correlations. There are obvious breed differences, denoting differences in culling reasons and culling policy. For example, milking speed seems more important in the Montbéliarde breed than for the two others. Longevity EBVs are always related to somatic cells score, female fertility and udder traits EBVs. Size and, to a lesser extent, feet and legs traits (at least as they are defined now) are poor predictors

of longevity. Note that if length of productive life is not adjusted for production, the observed correlations with milk yield or the aggregate index INEL are of 0.63 to 0.65 in the Montbéliarde and the Normande breeds. This correlation was not studied for the Holstein breed.

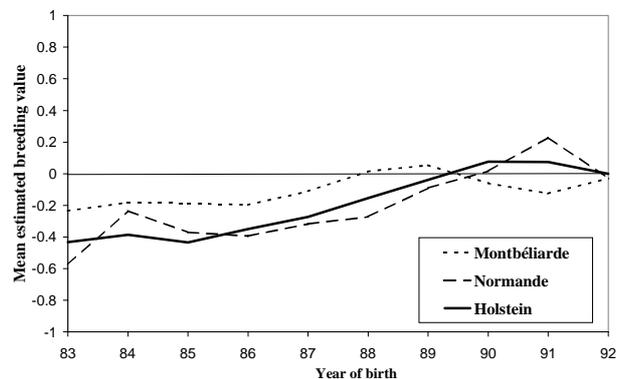
Mean EBVs for longevity are plotted in figure 5. For the three breeds, genetics trends have been favourable. The reasons of this optimistic result (given the known antagonism between for example fertility and milk yield) still need to be elucidated: it may reflect an incomplete adjustment for production traits. Strong and efficient type selection may also have played a role.

Table 2: Correlations between production-adjusted length of productive life EBV and EBVs for other traits.

EBV for:	Montbé- liarde	Normande	Holstein
INEL ^(a)	0.18	0.15	0.05
Milk yield	0.29	0.33	0.23
Protein content	-0.21	-0.26	-0.25
Cells score	0.29	0.29	0.34
Female fertility	0.33	0.28	0.26
Milking speed	0.33	0.15	0.17
Udder depth	0.25	0.33	0.39
Udder cleft	0.27	0.17	0.24
Fore udder	0.20	0.25	0.25
Rear udder	0.27	0.28	0.24
Height at sacrum	0.01	0.10	0.07
Feet and legs	0.15	0.23	
Rear leg set	-0.09	0.14	-0.01
Muscularity	-0.09	0.05	

^(a) = Protein yield (kg) + 0.3 Protein content (%)

Figure 5 Genetic trends for the 3 major breeds (in genetic standard deviation).



4.3 Reliabilities

Table 3 presents the fraction of bulls with an approximate reliability above 50, 60 or 70% after the June 98 and November 98 runs, in the Holstein breeds. This table exemplifies the main drawback of the current evaluation : it arrives too late. Most young

bulls get their first longevity proof (at least) one year after their EBVs for production and type traits is known. The proof is virtually useless for selection decisions. Two directions are foreseen to partly circumvent or reduce this problem: a better recording system, with a proper knowledge of true culling dates should reduce the lag that has to be imposed to properly define a record as uncensored (see above). Of course, inclusion of information from early predictors (type traits, somatic cell score, female fertility) is also envisioned.

Table 3: Fraction of Holstein AI bulls with reliability above a given level, depending on their year of birth, in November 1998 (in parentheses: June 1998)

Bulls born in	% with a reliability above		
	0.50	0.60	0.70
90	100 (99)	97 (87)	66 (37)
91	91 (72)	65 (33)	17 (4)
92	38 (16)	11 (2)	

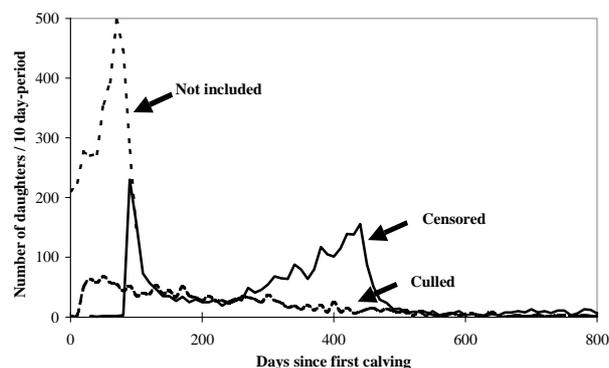
4.4 A recurrent problem

Run after run, a problem that was not anticipated was encountered: EBVs of very good young bulls (for production and type traits) are suddenly going up, frequently by more than one genetic standard deviation, when these bulls get a large batch of second crop daughters. They later come back to values more consistent with their starting point. A typical example (among quite a few others) is Cantadou, a Montbéliarde bull whose longevity evaluation went from +1.1 (just after getting his first second-crop daughters), to +1.6, +2.8 (1882 daughters) and +0.7 (7140 daughters). Needless to say, such changes are not well perceived by the breeders! One possible interpretation is the following: Cantadou is a good bull for production, type and somatic cell count. This is well known and when a farmer gets a second crop daughter, there is usually no reason to get rid of the cow *during* her lactation. Most cullings are at the end of lactation, for example linked to fertility problems. Calving is rather seasonal in France and when the evaluation of such a bull includes a massive number of new daughters, most of them censored in the middle of their first lactation, this gives a big push to their EBVs. It is only when most of them have reached the end of their first lactation that inferences on the impact of their sire on culling risk can be more safely drawn (see figure 6b for Cantadou). A first tentative to reduce this problem was implemented for the June 98 evaluation by discarding all cows that did not have the opportunity to live at least two months before τ_{max} . Figure 6a displays the impact of such rule for Cantadou for the June 98 evaluation. This is probably not enough. It is envisioned to increase to 5

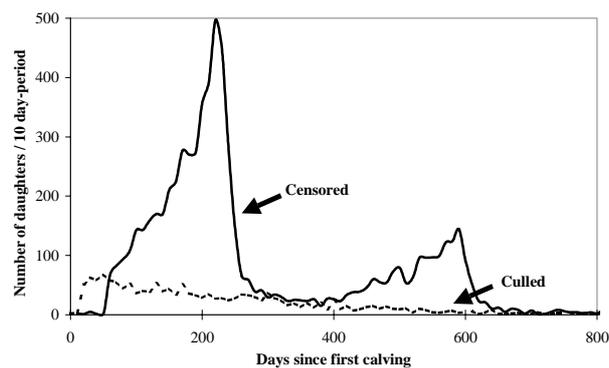
months this period (i.e., cows must have had the opportunity to live at least 5 months before being included in the analysis). This is somewhat similar to exclusion of extremely short lactations in progress for production evaluations. It has no real impact on the (small) reliabilities of young sires and should increase stability of proofs and therefore, enhance credibility.

Figure 6: Distribution of the records of daughters of the Montbéliarde bull “Cantadou” per period of 10 days

a) June 1998 evaluation



b) November 1998 evaluation



Conclusion

Length of productive life is a complex trait. Culling policy is influenced by many factors that are continuously changing over time. An attempt to account for all these changes necessarily results in a complicated model and a long and costly evaluation. The number and the importance of the assumptions made here are reduced compared to other approaches. All the available information is properly used. Strong genetic differences between bulls are detected. As expected, they are associated with several other functional traits, but they do not need a specific data recording scheme or an appropriate weighing of these functional traits. The main limitation of the approach is inherent to the trait itself: information on ability to delay culling is rather limited during early life and this leads to evaluations that are too inaccurate to be used in selection programs when selection decisions are made. There is a imperious necessity to enrich this information using relevant early predictors.

References

- Ducrocq, V., and Casella, G., 1996. A Bayesian analysis of mixed survival models. *Genet. Sel. Evol.*, 28: 505-529.
- Ducrocq, V., and Sölkner, J. 1998a. Implementation of a routine breeding value evaluation for longevity of dairy cows using survival analysis techniques. *Proc. 6th World Congr. on Genet. Appl. to Livest. Prod.*, Armidale, Australia, 23: 359-362.
- Ducrocq, V., and Sölkner, J. 1998b. «The Survival Kit V3.0», a package for large analyses of survival data. *Proc. 6th World Congr. on Genet. Appl. to Livest. Prod.*, Armidale, Australia 27: 447-448.
- Kalbfleisch, J.D. and Prentice, R.L. 1980. *The statistical analysis of failure time data*. Wiley, New York, NY.

Appendix: approximate heritability on the original scale

The model on the logarithmic scale can be written (Ducrocq and Casella, 1996):

$$\log(T_m) = \frac{1}{\rho} \mathbf{x}'_m \boldsymbol{\beta} + \frac{1}{\rho} s_q + \frac{1}{\rho} \omega_m \quad [A1]$$

where ω_m follows an extreme value distribution, with variance $\pi^2/6$. Therefore:

$$\begin{aligned} T_m &= \exp\left\{\frac{1}{\rho} \mathbf{x}'_m \boldsymbol{\beta}\right\} \exp\left\{\frac{1}{\rho} s_q + \frac{1}{\rho} \omega_m\right\} \\ &= K_i \exp\left\{\frac{1}{\rho} s_q + \frac{1}{\rho} \omega_m\right\} \end{aligned} \quad [A2]$$

On the observed scale (scale of T_m), in a particular "environment" characterised by a specific value of $\mathbf{x}'_m \boldsymbol{\beta}$, the heritability can be defined as 4 times the sire variance divided by the total variance after correction for systematic environmental factors, i.e., $h^2_0 = 4 A / B$ with:

$$A = \text{Var}\left[K_i \exp\left\{\frac{1}{\rho} s_q\right\}\right] \quad [A3]$$

$$B = \text{Var}\left[K_i \exp\left\{\frac{1}{\rho} s_q + \frac{1}{\rho} \omega_m\right\}\right] \quad [A4]$$

To calculate these expressions (deleting K_i , which, for a given environment, cancels out in the formula for h^2_0), it is necessary to compute the variance of a

function $g(\cdot)$ of the random variables s_q and ω_m . For this purpose, we will use the *delta method* which is based on a Taylor series expansion of $g(\cdot)$ around the mean of the random variable. The delta method specifies that, for a function $g(x)$ of a random variable x with expectation $\mu = E(x)$ and variance $\text{Var}(x)$, we have:

$$\text{Var}[g(x)] = [g'(\mu)]^2 \text{Var}(x) \quad [A5]$$

where $g'(\mu)$ is the first derivative of $g(\cdot)$ with respect to x , evaluated at $x = \mu$.

In the case of [A3], $g(s) = \exp\left\{\frac{1}{\rho} s\right\}$ and $\mu = E(s) =$

0. The first derivative of $g(\cdot)$ is:

$$g'(s) = \frac{1}{\rho} \exp\left\{\frac{1}{\rho} s\right\}$$

and therefore:

$$A = \left[\frac{1}{\rho} \exp\{0\}\right]^2 \text{Var}(s) = \frac{1}{\rho^2} \text{Var}(s) \quad [A6]$$

For [A4], let $u_m = \frac{1}{\rho} s_q + \frac{1}{\rho} \omega_m$. We take $g(u_m) = \exp(u_m)$. Then, it follows that:

$$\begin{aligned} \mu &= E(u_m) = E\left[\frac{1}{\rho} s_q + \frac{1}{\rho} \omega_m\right] \\ &= \frac{1}{\rho} 0 + \frac{1}{\rho} E(\omega_m) = \frac{1}{\rho} v \end{aligned} \quad [A7]$$

where v is the expectation of an extreme value distribution: $v = -\text{Euler's constant} = -0.5772\dots$

This leads to:

$$\begin{aligned} B &= \left[\exp\left\{\frac{1}{\rho} v\right\}\right]^2 \text{Var}(u_m) \\ &= \left[\exp\left\{\frac{1}{\rho} v\right\}\right]^2 \frac{1}{\rho^2} \left[\text{Var}(s) + \frac{\pi^2}{6}\right] \end{aligned} \quad [A8]$$

Hence, after simplification:

$$h^2_0 = \frac{4 \text{Var}(s)}{\left[\exp\left\{\frac{1}{\rho} v\right\}\right]^2 \left[\text{Var}(s) + \frac{\pi^2}{6}\right]} \quad [A9]$$

or:

$$h^2_0 = \left[\exp\left\{\frac{1}{\rho} v\right\}\right]^{-2} h^2_{\log} \quad [A10]$$

