

## BENEFITS FROM ANIMAL MODEL EVALUATION OF DAIRY CATTLE IN FRANCE

B. BONAITI and D. BOICHARD  
Institut National de la Recherche Agronomique,  
Station de Génétique Quantitative et Appliquée,  
78350 JOUY-en-JOSAS, FRANCE

### SUMMARY

This paper presents a concrete application of an Animal Model genetic evaluation. Such a model is becoming the international reference for evaluation purposes. After a description of the technical assumptions, emphasis is on problems encountered during the implementation. The correlation with previous estimates is high, above 0.975 for AI bulls, above 0.90 for other males. The genetic trend estimated with the animal model remains unchanged in closed populations (*Normande* and *Montbéliarde*). However, in the *Friesian* breed, which is being absorbed into the North American *Holstein* strain, it appears 23 % lower and different points seem to show that it is underestimated. Finally, advantages and limits of the animal model are presented. The main advantage is its theoretical clarity and, consequently, its ease to be explained. The main limit may be the excessive confidence of breeders who sometimes tend to forget that results will always remain highly dependent on the quality of the data.

### INTRODUCTION

BLUP applied to an animal model (AM) is now the international reference for genetic evaluation. After the USA (Wiggans *et al.*, 1988 ; Wiggans and Van Raden, 1989) and Canada, a lot of countries are working on its practical implementation for dairy cattle. In France, the new evaluation system replaced in March 1990 the old one, derived from the MCC. In this paper, we present the consequences of this change, with emphasis on the technical assumptions, the problems encountered during the implementation and the advantages of the new system.

### TECHNICAL ASSUMPTIONS FOR THE ANIMAL MODEL APPLICATION IN FRANCE

#### 1. Choice of a model of analysis

The situation of dairy production in France, which is comparable to many other European countries, makes necessary the choice of a rather complex model of analysis. Because the herd size is small (about 30 cows on average), several lactations for each cow are taken into account. Moreover, with a model including only one fixed effect, *i.e.* a herd-year-season-parity effect, as in the USA, a lot of data would have been lost because the elementary cell would often be too small. On the other hand, the *Friesian* breed (70 % of the French dairy cows) is being absorbed into the north-american *Holstein* strain. This evolution involves big changes in management system, particularly in age at first calving. Prior correction for age would be difficult to perform, because age at calving is not independent on genetic level. Indeed, a drastic simplification of the model used in our previous evaluation

method (IF2) for the sake of being able to implement the animal model is deemed undesirable. Finally, the model used is the following:

$$Y_{ijklmnr} = U_i + P_i + HY_{jk} + LN_{klr} + M_{klrm} + A_{klrn} + E_{ijklmnr}$$

with  $U_i$  : genetic value of cow  $i$ ,

$P_i$  : permanent effect of cow  $i$ ,

$HY_{jk}$  : effect of herd  $j$  during year  $k$ ,

$LN_{klr}$  : effect of parity  $l$ , within year  $k$  and region  $r$ ,

$A_{klrn}$  : effect of class of age  $n$  at calving (for first lactation) or effect of class of calving interval  $n$  (for lactations 2 or 3), within year  $k$ , parity  $l$  and region  $r$ ,

$M_{klrm}$  : effect of month  $m$  of calving, within year  $k$ , parity  $l$  (lactation 2 and 3 are grouped) and region  $r$ ,

$E_{ijklmnr}$  : residual effect.

$U$ ,  $P$  and  $E$  are assumed to be normally distributed with expectation  $Qg$ , 0 and 0, and variance  $A\sigma_a^2$ ,  $I\sigma_p^2$  and  $W\sigma_e^2$  respectively.

The heritability and repeatability are assumed to be equal to 0.3 and 0.5 respectively for yields, and 0.5 and 0.7 for contents.  $W^{-1}$  is a diagonal matrix weighting each performance, assuming a weight of 1 in first lactation and 0.8 later. The weights are reduced for extended lactations. Groups are defined for unknown parents only (Westell *et al.*, 1988), according to sex, birth year and origin of the progeny. In most cases, both parents are unknown and put in the same group. In the *Friesian* breed, because of the large heterogeneity of origins, the number of groups is rather high and reaches 400.

Single trait analysis are performed quarterly, within breed. Results are expressed relative to two rolling bases, different for males and females and updated each year.

## 2. Definition of the data set

Before the analysis, in-progress lactations are extended, as in Mocquot (1978). Milk, fat and protein yield are adjusted for parity and lactation length (LL, in days) with the following multiplicative coefficient:

$$C \cdot 385 / (LL + 80)$$

with  $C$  equal to 1.3, 1.17 and 1.1 in lactations 1, 2 and 3 respectively.

The evaluation is divided into two steps. The first one, the most important one, is solution of the animal model equations, *sensu stricto*. All animals with a first lactation started after September 1st, 1974 and all their pedigree information, are included. The first three lactations are considered only. A second or third lactation is accepted only if the first one is known and performed in the same herd. Table 1 shows the size of the data set corresponding to the three main breeds.

As a second step, the other lactations (rank higher than 3, lactations when the first one is unknown or performed in another herd) are taken into account for cow evaluation. The same strategy is used as in Wiggins *et al.* (1988). However, this additional information is not used for the evaluation of the relatives.

In the *Friesian* breed, the US or Canadian breeding value of the foreign bulls, which semen is imported, is included in the analysis according to the following methodology. First breeding values are converted on the French scale with a conversion formulae (here, only the  $b$

Table 1. Description of the data set analysed with the animal model in January 1990 for the three main breeds (1)

	Breeds		
	<i>Friesian</i>	<i>Normande</i>	<i>Montbeliarde</i>
Number of females	7667160	2084331	1242758
Number of females with records (2)	4623997	1388779	887465
Number of records (milk)	10106199	2909448	1939389
Number of males	131559	41664	39211
Number of AI males	7845	3587	1671
Number of herds	110634	50440	22619
Number of years	16	16	16
Number of herd-year	755737	239594	144324
Number of regions	8	7	3
Number of fixed effects other than herd-year	5888	5152	2208
Number of groups	400	75	35
Number of dam-daughter pairs in different herds	902520	150686	107043

(1) 12 other small breeds were also analysed, with between 1,000 and 80,000 cows each. The method will also be extended to dairy sheep and goat data.

(2) Records of *Friesian* cows without pedigree were discarded.

coefficient has an impact). Fictive daughters are then generated, with one record each, equal to twice the deregressed breeding value of the bull. The number of daughters considered is a function of the repeatability of the foreign evaluation, but is limited to 100. These females are grouped in a fictive herd, whose effect is fixed to zero. Their dams, assumed to be unknown, are put in a particular group, whose estimate plays the role of the  $a$  coefficient of the conversion formulae. Furthermore, when the foreign breeding value of a bull is considered, his pedigree is not taken into account, to avoid redundant information.

### 3. Solution strategy

Since a supercomputer is not available to us, BLUP equations are to be solved without simultaneously storing all records and all estimates in core. The method used has been described in Ducrocq *et al.* (1989). Each iteration implies reading a data set sorted by region and herd, and two other data sets including respectively breeding values and environmental effects estimates of the previous iteration. The main principles, described in the appendix, are based on the 4 following steps :

1) absorption of permanent effects.  
2) pseudo-absorption of breeding values, assuming the breeding values of relatives to be known and using their last estimates.

3) bloc-iterative solution of 5 groups of equations : fixed effects other than herd-year, herd-year effects, female breeding values, male breeding values and group effects. This requires a process to exchange breeding values between a cow and its relatives in different herds. From this point of view, this method is derived from the previous French evaluation system (Poutou *et al.*, 1981 ; Bonaiti *et al.*, 1984) and the algorithm proposed by Bonaiti and Briand (1986).

4) Direct solution of a subsystem of equations involving breeding value of animals with only one parent considered (Poivey, 1986). This method is applied twice : first, to females and their dams within herd, assuming that males and females in others herds had known breeding value ; second, to males and all their progeny, assuming breeding values of the bull dams and bull mates were known.

The whole *Friesian* application requires 6 megabytes of storage in core. External storage on disk reaches 160 bytes per animal, including pedigree information, records, breeding values, work files, information exchanged between herds, left-hand side and estimates of fixed effects. For each million of animals, one iteration requires 3 minutes of CPU and 400 megabytes of input/output, each file being read or written 2 or 3 times.

### 4. Differences with the previous evaluation system ("IF2")

From 1978 to 1989, breeding values were obtained using a method comparable to the MCC, but already quite close to an animal model approach. Sire breeding values were obtained from progeny records adjusted for their dams breeding values. Cow breeding values were estimated from pedigree and own performance information, but not from progeny. Therefore, as in the animal model, male breeding values were not biased by differences in genetic levels of mates. But the models for males and females were not exactly the same.

The main differences appeared in the computing strategy. While the animal model application is entirely supported by the theory, it was not true for the IF2 method. A special effect, accounting for culling of the cows, avoided the definition of an individual effect, but had no clear mathematical justification. At each run, only one iteration was performed. Moreover, breeding values of the cows with no new record during the last three years, were assumed to be known and were never updated. The same strategy was applied to environmental effects more than 3 year old.

Another main difference with IF2 involves the analysed traits. In the old system, protein and fat were combined together and three traits were evaluated : milk yield, "useful" yield and "useful" contents. But, as the importance of protein is increasing in France relatively to fat and, on the other hand, as the comparison with foreign results is much easier when the evaluated traits are the same, fat and protein are analysed separately in the new system.

## PROBLEMS ENCOUNTERED DURING THE ANIMAL MODEL IMPLEMENTATION

### 1. Comparison with previous results

Results for milk yield with both methods may be directly compared. For the other traits, AM fat and protein breeding values are combined before the comparison with useful yield and contents (fat + 1.2 Protein) obtained with IF2.

For AI bulls, results for both methods are in virtually complete agreement. The within year variabilities of breeding value are the same. The correlation between AM and IF2 breeding values reaches at least 0.975 for yields and 0.99 for contents in all breeds.

For females, the results presented here correspond to breeding values obtained from the first step of the AM application, without reintegrating excluded records. This may explain why the correlation is only moderate, between 0.75 and 0.90, and why it increases with the quantity of information included in the AM evaluation. For the same reason, and also because of the repeatability of the breeding values, involved in the IF2 computations, used to be overestimated, the variability of AM breeding values is 10 % lower than with the previous method for females.

### 2. Estimated genetic trends

In the closed populations (*Normande* and *Montbéliarde*), the old and the new estimates are in perfect agreement. However, for the *Friesian* breed, which is being absorbed into the *Holstein* strain (figure 1), the new estimate (1280 kg milk between 1970 and 1982), based on AI bulls, appears 23 % lower than the previous one (1660 kg).

At the beginning of the 80's, *Holstein* bulls were tested with F1 crossbred progeny and, because of heterosis (125 kg, according to Van Der Werf and De Boer, 1989), were probably overestimated by about 250 kg. But now, cows are 3/4 or 7/8 *Holstein*, and heterosis tends to disappear. A retrospective analysis may be closer to the true value than IF2, since this latter method did not analyse simultaneously all the data and consider old breeding values estimates as fixed. However, this first fact may explain only part, maybe half, of the surprising gap between both estimates.

The animal model estimate appears rather insensitive to the numerous hypothesis tested successively, as an increase of heritability, change in grouping strategy, inclusion of first lactations only or of the higher-producing part of the population only. The animal model estimate seems even higher than that obtained with a sire model. However, discarding records of cows without pedigree lead to a significative increase (+100 kg) in estimated genetic gain. This result is interesting to discuss. The cows without pedigree have probably not the same expected breeding value between herds within region and year. An unknown sire is likely to belong to the major strain (*European Friesian* or *Holstein*) used in the herd of her daughter. Discarding records of cows without pedigree partially destroys the false connection between herds through groups of unknown parents. It is clear now that attention should be paid to the definition of groups according to true genetic differences between herds.

Three facts lead us to believe that the genetic trend estimated by the animal model was

Figure 1. Genetic trend for milk yield in France and USA, estimated from AI Holstein bulls, according to the method of evaluation

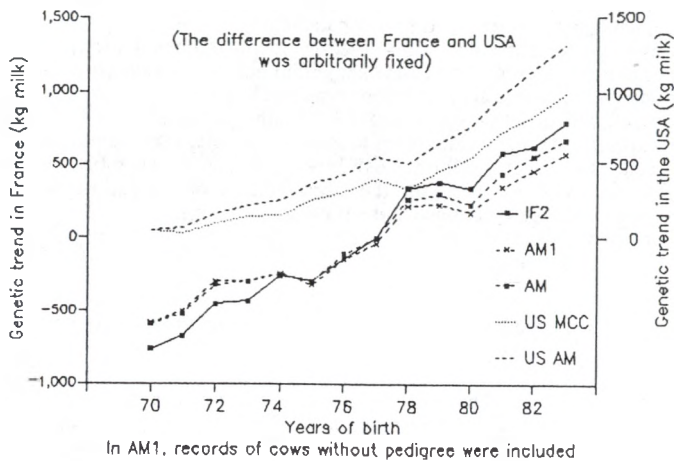
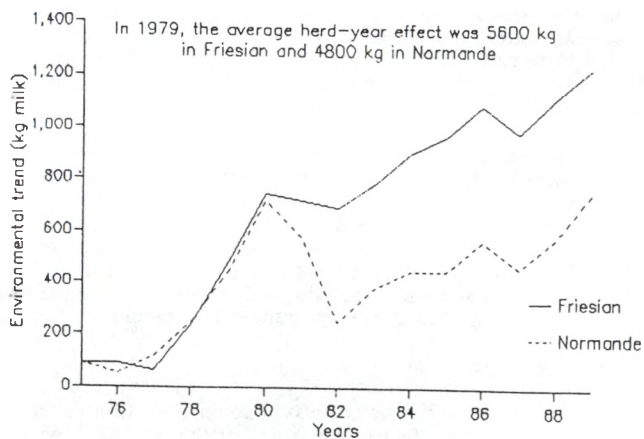


Figure 2. Trend of herd year effect in the 10,000 herds with both Friesian and Normande breeds





underestimated.

- More than 10,000 herds have simultaneously *Normande* and *Friesian* cows. Their herd-year effects (figure 2) are clearly divergent since 1980 (when *Holstein* progeny testing was generalised), while they are expected to stay parallel.

- The average daughter deviation (average record adjusted for fixed effects and breeding value of mate) does not remain constant within bull but is increasing with time. This means that a young bull is underestimated when compared to an old one.

- within sire and maternal grand sire, the breeding value of sons is decreasing with time, while it is expected to remain constant or to increase slightly, because of a better choice of dams. As a consequence, conversion formulae of breeding values between France and USA are less accurate than with IF2. Furthermore, the large increase of US estimated genetic gain (figure 1) contributes to a drastic decrease of the *b* values.

The problem remains still open.

## ADVANTAGES AND LIMITS OF ANIMAL MODEL

The properties of the AM have already been presented from a theoretical point of view or through numerical examples by Kennedy *et al.* (1988), Henderson (1988), Westell *et al.* (1988), and Ducrocq (1990). Theoretically, by using the matrix describing all relationships from the initial base population to the present one, the AM takes into account the evolution of genetic variability due to drift, selection and even partly to inbreeding. It is free of some assumptions -which are never fulfilled- implicit in a sire or sire + maternal grand sire models : assuming equal genetic level of mates (sire model) or dam of mates (sire + maternal grand sire) is no longer necessary.

### 1. The AM uses all the available information

Theoretically, assuming the variance components to be known, BLUP applied to the animal model provides the most likely values for all parameters. A retrospective analysis of all the data improves the estimated breeding values of previous generations and consequently provides the most accurate estimate of genetic trend. In contrast, an analysis restricted to the most recent data, as it was the case with IF2, may provide biased estimates of the genetic trend. However, as we saw above, this property of AM is highly dependent of a right description of the real genetic situation (model of analysis, grouping strategy).

Including all relationships gives the best estimation of cow breeding values, by estimating genetic differences between herds. Part of these differences, due to a different choice of sires, was already well estimated in simplified models, but the other part, due to cow selection, is better estimated using an animal model, which takes into account the within herd selection and connection from relatives in different herds. At the present time, we do not get all the profit related to this because of the rather few exchanges between herds (10%). However, from this point of view, an animal model evaluation will easily integrate the evolution of reproduction techniques, as the extension of embryo transfer for example.

### 2. The animal model is clear

The main disadvantage of the IF2 method was its complexity. Consequently, breeders encountered difficulties to understand the method as well as the results. However, a BLUP applied to a herd-year + sire (or sire + maternal grand sire) model, easier to compute and to explain, was voluntarily not chosen in France in 1978, mainly because it would have been too far from the biological reality (large variability of management due to *Holstein* absorption, small

herd size...). In addition, introducing several lactations for each cow, as required with small herd populations, would have substantially increased the computing cost of BLUP. Animal model is much easier to explain : the variance components, the definition of the model and the strategy of grouping fully describe the evaluation system. Discussion of these three points may be opened to all the breeders and is no longer restricted to few specialists.

In the model, each parameter has a biological meaning. Environmental effect estimates may be used by extension services, in order to analyse differences between herds for example. Providing both breeding values (U) and permanent effects (P) will help the breeders to discriminate between transmitting and producing abilities, *i.e.* between breeding value and future profitability. The producing ability (U+P) may be used for culling decisions.

Sires and dams, now evaluated on the same scale, can be directly compared to each other. Consequently, pedigree breeding values are more accurate and homogeneous, and the choice between embryos, females or semen to buy becomes easier.

A breeding value can be easily splitted into three components : a pedigree value  $U_a$ , a progeny value  $U_p$ , and an average record  $Y_c$  adjusted for all environmental effects (including permanent effect) :

$$U = (w_1 U_a + w_2 U_p + w_3 Y_c) / (w_1 + w_2 + w_3)$$

where the three weights are derived from the amount of pedigree and progeny information on the animal. Publication of these three components will help breeders to understand breeding values.

Lastly, the animal model, because of its clarity and its simplicity, has become the official method and the international reference. The use of the same method all over the world will make breeders more confident in breeding values, so selection should be more efficient. Collaboration between countries will be easier, allowing a better definition of models and exchanges of computing algorithms. The first instance of such a collaboration is the decision by the Interbull committee to ask its members to make available sire average deviations for computation of conversion factors.

### 3. Traits considered

A particularity of the French situation is the change in the traits considered. Although protein was becoming more and more important relative to fat for milk industry, the old system was not directed towards the supply of breeding values for fat and protein separately. From the farmer's point of view, fat and protein breeding values are now available and this may be among the largest consequences of the animal model implementation.

### 4. Limits of the animal model

The fame of animal model evaluation is becoming so large that the resulting breeding value estimates are considered as true values by breeders, whatever the data they are based on. However, these estimates may be biased since environmental factors are not fully controlled. It is important to remember that good results always depend on a good method applied to a good data set. Let us recall the main potential problems.

First, cows from a given region are often inseminated by technicians from a single AI center, with semen of bulls belonging to that same AI center. In such a situation, the connection between bulls of different AI studs is only due to pedigrees. Only those breeders who use semen from several centers may increase the connection which, even then, remains often small. Progeny testing from different organisations in the same region would be an efficient,



sure and thus unquestionable way to compare bulls among AI studs. This direct comparison is always better than the indirect one, based on service bulls or on common pedigrees, through the relationship matrix. In France, it has been decided that each of the five AI organisations should produce at least ten daughters issued from twenty bulls in another region than its original one. The first daughters of this programme will have records in fall 1990. This collaboration will be increased in order to detect possible biases between selection organisations.

Second, preferential treatments may bias the estimated breeding values. It is not only true for cows, but also for progeny of males when their semen is very expensive. Because the use of "imported" semen (from foreign countries or other regions) is often associated with preferential treatment, estimates of genetic differences between countries or regions, mainly based on these data, may be biased, the importing regions being underestimated. The use of BST -legally or not- oriented towards the generation of very high breeding values, may drastically increase these biases. Actions which would reduce its effect, such as prior elimination of records, detection of BST treatments, use of other traits like part of lactation or within herd rank, must be undertaken.

Third, practical experience showed that group definition may influence results. Indeed, the expected breeding value, assumed to be constant within group, may vary according to region, birth year, or even production level of herd. Thus a correct application of the animal model assumes 1) the use of the largest possible number of pedigrees, 2) the definition of groups as accurate as possible, according to the real genetic level of the animals.

As a conclusion, the animal model has beautiful theoretical properties which are a great progress in the evaluation field. But to keep them in the applications, extensive attention is to be paid to the quality of the data and to the choice of the model of analysis.

#### ACKNOWLEDGMENTS

We are grateful to G.R. Wiggans for his constant interest and his useful comments provided all along our work on the animal model implementation.

#### REFERENCES

- BONAITI B., BRIEND M., 1986. *Génét. Sél. Evol.*, 18, 41.  
BONAITI B., MOCQUOT J.C., POUTOUS M., 1984. IDF-EAAP Symposium, Praha, 09/14-16/1984  
DUCROCQ V., 1990. *Prod. Anim. INRA*, 5 (in press)  
DUCROCQ V., BOICHARD D., BONAITI B., BARBAT A., BRIEND M., 1990. *J. Dairy Sci.*, 73 (in press)  
HENDERSON C.R., 1988. *J. Dairy Sci.*, 71 (suppl2), 1.  
KENNEDY B.W., SCHAEFFER L.R., SORENSEN D.A., 1988. *J. Dairy Sci.*, 71 (suppl2), 17.  
MOCQUOT J.C., 1978. *Ann Génét. Sél. Anim.*, 10,359.  
POIVEY J.P., 1986. *Génét. Sél. Evol.*, 18,321.  
POUTOUS M., BRIEND M., CALOMITI S., DOAN D., FELGINES C., STEIER G., 1981. *Bull. Tech. Inf.*, 361, 433.  
VAN DER WERF J.H.J., DE BOER W., 1989. *J. Dairy Sci.*, 72, 2615.  
WESTELL R.A., QUAAS R.L., VAN VLECK L.D., 1988. *J. Dairy Sci.*, 71, 1310.  
WIGGANS G.R., MISZDAL I., VAN VLECK L.D., 1988. *J. Dairy Sci.*, 71 (suppl2), 54.  
WIGGANS G.R., VAN RADEN P.M., 1989. USDA-DHIA report H2, 8p.

## APPENDIX

### DESCRIPTION OF THE ALGORITHM

Note H the herd-year effect, F the other fixed effects, U the genetic value, and P the permanent effect. A subscript k indicates the estimates at iteration k.

#### Preliminary step

Construction of the F x F and H x H left-hand side (LHS), after absorption of P equations and "pseudo-absorption" of U equations

Construction of the diagonal of the U x U LHS matrix, after absorption of P equations

#### Iterative procedure (iteration k)

##### *1. Estimation of $F_k$* (this part may be skipped)

Correction of U's right-hand side (RHS) for all relatives using  $U_{k-1}$

Correction of F's RHS for  $H_{k-1}$  effects

Absorption of P equations into F's and U's RHS

Pseudo absorption of U equations into F's RHS

Estimation of  $F_k$

##### *2. Estimation of $H_k$ and first estimation of U for females*

Absorption of P equations into H's and U's RHS

Correction of U's RHS for all the relatives ( $U_{k-1}$ )

Correction of U's and H's RHS for  $F_k$

Pseudo-absorption of U equations into H's RHS

Estimation of  $H_k$

Correction of U's RHS for  $H_k$  and for all relatives outside the herd

Estimation of  $U_k$  for females within herd, using Poivey's method

Construction of males RHS

. correction for bull dams and bull mates

. absorption of progeny equations

##### *3. Solution for males using Poivey's method*

Solution for males U's

Construction of groups RHS (contributions of males)

##### *4. Solving for females and groups*

Backsolution of females U's

Forwarding information to relatives in other herds

Construction of groups RHS (contributions of females)

Computation of group effects solutions

##### *5. Sort of the file containing the information exchanged between herds*