

ÉTUDES SUR LA PRODUCTION LAITIÈRE DES BOVINS

II. — INTÉRÊT DES LACTATIONS PARTIELLES POUR LA SÉLECTION.

a) ÉTUDE CRITIQUE DE DIFFÉRENTES MÉTHODES D'ESTIMATION DES PRODUCTIONS TOTALES A PARTIR DES RÉSULTATS DE LACTATION PARTIELLE

T. AURAN ⁽¹⁾ et J. C. MOCQUOT

*Station de Génétique quantitative et appliquée,
Centre national de Recherches zootechniques, I. N. R. A.,
78350 Jouy en Josas*

RÉSUMÉ

La prévision des productions laitières totales par lactation à partir des résultats partiels présente un intérêt pratique pour l'indexation rapide et précise des mâles sur descendance ainsi que pour un choix précoce des femelles intra-troupeau. Elle est possible compte tenu des liaisons statistiques importantes qui existent entre les productions aux différents stades de la lactation. Différentes méthodes d'extrapolation proposées jusqu'à maintenant sont décrites et analysées de façon critique.

Quelques auteurs ont essayé de formuler une équation de la courbe de lactation ; les expressions obtenues étant trop complexes, la plupart des autres méthodes effectivement utilisées ne font intervenir qu'un nombre limité de coefficients. Leur efficacité relative a été comparée à l'aide de deux critères :

- Le niveau de corrélation entre les productions estimées et réelles.
- La concordance de variance des estimées et des productions réelles.

Leur commodité d'utilisation pour une réalisation pratique des calculs est également discutée. La plus simple et la plus employée est la méthode multiplicative de la production partielle cumulée, le coefficient utilisé étant le rapport moyen des productions totales aux productions partielles de même durée (même nombre de contrôles mensuels). Elle tend à sous-estimer les mauvaises productions et à surestimer les meilleures. Elle conduit cependant à des estimées de variance proche de la variance réelle.

Les autres méthodes conduisent à des estimées par régression sur la production partielle cumulée (régression simple, régression modifiée, méthode P), ou sur chacun des contrôles mensuels (régression multiple). Elles sont plus précises en terme de corrélation entre estimées et valeurs réelles, mais la variance des estimées est nettement inférieure. La méthode par régression simple, (ou ses dérivées), est beaucoup plus facile à mettre en œuvre car elle n'implique le stockage que d'une seule variable. Elle est presque aussi précise que la régression multiple. La méthode P

⁽¹⁾ Adresse permanente : Institut de Génétique animale, Université agricole de Norvège, AS-NLH (Norvège).

qui tente de prévoir la production restant à réaliser au lieu de la production totale est intéressante dans la mesure où la meilleure variable prédictrice de cette production à estimer semble être le dernier contrôle de la lactation partielle et que cette variable est en général disponible en même temps.

Le calcul des régressions intra-troupeau accroît de 10 à 20 p. 100 la précision des estimations mais il implique l'utilisation des moyennes de troupeau pour chaque variable prédictrice.

L'analyse de différents facteurs de variation indique qu'il faut disposer de jeux distincts de coefficients au moins pour les différents numéros de lactation et différents caractères, mais aussi éventuellement pour les différentes races, saisons de vêlage et niveaux d'étable.

Enfin l'extrapolation des lactations « partielles » consécutives à l'élimination d'animaux en cours de lactation est délicate dans la mesure où la cause d'élimination est souvent inconnue. La courbe de lactation des femelles éliminées précocement pour insuffisance de production est en effet différente et implique l'utilisation de coefficients appropriés et difficiles à estimer.

INTRODUCTION

D'un point de vue pratique, la prévision des productions totales par lactation à partir des résultats partiels présente pour la sélection un double intérêt :

— Au niveau du troupeau elle peut permettre une estimation précoce de la valeur génétique des femelles pour une sélection plus efficace, au moment du vêlage notamment.

— Au niveau de la sélection sur descendance des mâles elle peut permettre une estimation plus rapide de leur valeur génétique sur l'ensemble des résultats enregistrés pour leurs filles en lactation, donc une estimation également plus précise. En effet, lorsque les calculs d'index sur descendance s'effectuent à date fixe à partir des seules lactations complètes, toute estimation précoce de la valeur d'un taureau ne fait intervenir qu'un nombre limité de ses filles, celles qui ont vêlé en moyenne à un âge plus réduit et ont eu pour la plupart des lactations courtes. Une telle situation conduit à des valeurs biaisées (VAN VLECK, 1962 ; MOCQUOT et POUTOUS, 1971) et peu précises alors qu'au même instant sont également connus et disponibles les résultats partiels de la quasi totalité de l'échantillon de filles. VAN VLECK et HENDERSON (1961 *g*) ont d'ailleurs montré que l'utilisation de 20 résultats partiels de lactations de 5 mois et de 10 résultats complets faisait passer la précision à 0,78 contre 0,61 pour les seules 10 lactations complètes, et que l'estimation obtenue était presque aussi précise que celle calculée à partir des 30 lactations complètes ($R^2_{\hat{g}\hat{g}} = 0,80$).

L'efficacité d'une telle prévision statistique de la lactation totale à partir d'un nombre réduit de contrôles partiels est imputable aux fortes corrélations existant entre les mesures de production laitière effectuées aux différents stades de la lactation. Étudiées à de nombreuses reprises, ces liaisons sont citées par HOFFMANN dès 1959 et plus tard discutées, pour les 8 plus importantes études portant sur un effectif global de 37 963 lactations, par GUSTAFSON (1972). Le calcul, à partir de ces chiffres, de valeurs moyennes pondérées conduit à estimer les coefficients de corrélation phénotypique entre la production totale par lactation et les productions journalières enregistrées au cours des différents contrôles laitiers mensuels respectivement à 0,60 ; 0,76 ; 0,82 ; 0,85 ; 0,86 ; 0,86 ; 0,84 ; 0,80 ; 0,71 et 0,58. Les

coefficients de corrélation entre les productions cumulées correspondantes et les productions totales s'établissent ainsi à : 0,66 ; 0,79 ; 0,84 ; 0,89 ; 0,92 ; 0,94 ; 0,96 ; 0,98 et 1,00. Une étude réalisée par KEOWN et VAN VLECK (1971), portant sur 63 300 lactations, conduit à des valeurs légèrement supérieures, ces auteurs constatent également les liaisons les plus étroites entre le 4^e et le 7^e contrôle. La même tendance peut être observée pour les corrélations génétiques estimées par SEARLE (1961 *b*), VAN VLECK et HENDERSON (1961 *a*) ainsi que par KEOWN et VAN VLECK (1971) et dont les valeurs semblent plus élevées que les liaisons phénotypiques.

L'utilisation de ces relations pour l'estimation précoce de la valeur génétique des reproducteurs mâles sur descendance semble donc justifiée. Avant d'en aborder l'étude détaillée dans les conditions françaises il nous a semblé utile de faire le point des différentes méthodes d'extrapolation des lactations utilisées jusqu'alors, de les discuter et si possible d'en tirer des enseignements quant aux influences de différents facteurs susceptibles d'en modifier les paramètres et quant aux méthodes à mettre en œuvre pour en tenir compte.

I. — DESCRIPTION DES MÉTHODES D'EXTRAPOLATION DES LACTATIONS PARTIELLES

1. — *Utilisation d'une courbe de lactation standard*

Préalablement à la présentation des méthodes d'extrapolation proprement dites il est important de signaler un certain nombre d'études visant à prévoir la production laitière à chaque stade de lactation en formulant de façon algébrique la courbe de lactation. De telles équations des courbes de lactation ont été proposées par GAINES (1927), VUJICIC et BACIC (1961), NELDER (1966) et WOOD (1967). La plupart sont soit des polynômes de degré élevé soit des fonctions exponentielles faisant intervenir des paramètres largement influencés par les différents facteurs de variation de la production laitière. Ceci réduit considérablement leur possibilité d'utilisation pratique pour une simple estimation des productions totales. WOOD (1967) propose cependant des facteurs multiplicatifs dérivés de son expression algébrique pour l'estimation des productions totales des filles des taureaux de testage. La concordance entre les index ainsi estimés et les index réels est discutée dans le rapport 71-72 du MMB.

Une autre méthode de prévision, sans être une véritable méthode d'extrapolation est proposée par MOXLEY (1968). Elle consiste à exprimer la production individuelle à chaque contrôle en production équivalente à 140 jours de lactation pris comme stade de référence compte tenu d'une courbe de lactation standard. Ces valeurs sont ensuite corrigées par régression intra-troupeau sur le stade de lactation afin de tenir compte de l'effet troupeau-saison sur la courbe de lactation. Cette méthode nous semble digne d'intérêt dans la mesure où elle permet de classer simplement chaque mois les vaches du troupeau par rapport à une moyenne d'étable très contemporaine. Il nous semble d'ailleurs possible d'en dériver une formule de prévision de la lactation totale de la forme :

$$\widehat{Y} = A \times \Sigma C_{110}$$

où les C_{140} représenteraient les productions aux différents contrôles, corrigés pour l'âge, le stade de lactation, le troupeau et exprimés à un stade de référence de 140 jours de lactation ;
 et A un facteur d'extrapolation fonction du nombre de contrôles connus.

2. — Utilisation de facteurs multiplicatifs

Du fait de l'existence généralisée de contrôles laitiers mensuels la méthode d'extrapolation la plus employée consiste à multiplier chaque contrôle, ou chaque production cumulée, par un facteur constant compte tenu du numéro d'ordre de ce contrôle, ou du stade de la lactation partielle,

$$\widehat{Y} = C \cdot Y_p \quad (1)$$

De tels coefficients d'extrapolation ont été estimés dès 1942 par CANNON *et al.* De nombreux auteurs ont ensuite calculé ces facteurs aussi bien pour des productions partielles cumulées que pour des résultats de contrôle mensuel, citons : MADDEN *et al.* (1955, 1956, 1959), LAMB et MCGILLIARD (1960 *a*, 1967 *a*), SYRSTAD (1964), McDANIEL *et al.* (1967), APPLEMAN *et al.* (1969) ; WOOD (1969) ; BAPTIST (1972) et enfin KEOWN et VAN VLECK (1973).

Certains auteurs comparent cette méthode d'extrapolation avec d'autres : MILLER *et al.* (1972 *a* et *b*) ou étudient les facteurs de variation de ces coefficients sans en donner la valeur : VAN VLECK et HENDERSON (1961 *c*), MILLER *et al.* (1967). Dans tous les cas les coefficients sont estimés à partir de productions totales et partielles d'échantillons aussi les valeurs obtenues varient elles selon les données, les races, les hypothèses ou les éventuelles corrections appliquées à ces données. Le calcul de valeurs moyennes pondérées des différents chiffres de la littérature conduit aux valeurs suivantes, pour 19 432 premières lactations, compte tenu du numéro d'ordre du dernier contrôle de la lactation partielle : 8,51 ; 4,17 ; 2,84 ; 2,19 ; 1,79 ; 1,53 ; 1,34 ; 1,20 et 1,10.

3. — Estimations par régression

D'après MADDEN *et al.* (1959), ce type d'estimation de la lactation totale a été introduit par HARVEY en 1956. Différentes équations de régression ont ensuite, été proposées par MADDEN *et al.* (1959), VAN VLECK et HENDERSON (1961 *b*, *d*, *f*) APPLEMAN *et al.* (1969), BAPTIST (1972), MILLER *et al.* (1972 *a*, *b*) enfin par SCHLOTE (1973). La plus souvent utilisée et la plus simple consiste en une *régression simple* sur la production partielle :

$$\widehat{Y} = \bar{Y} + b(Y_p - \bar{Y}_p) \quad (2)$$

où \bar{Y} et \bar{Y}_p représentent par exemple les moyennes intra troupeau des productions totales et partielles. VAN VLECK et HENDERSON (1961 *b*, *d*) ainsi que MILLER *et al.* (1972 *b*) ont également tenté d'estimer la lactation totale par *régression multiple* intra-troupeau sur les productions aux différents contrôles mensuels de la lactation partielle :

$$\widehat{Y} = \bar{Y} + \sum_{i=1}^m b_i(Y_{pi} - \bar{Y}_{pi}) \tag{3}$$

où i est l'indice du contrôle considéré et Y_{pi} la production laitière correspondante, de moyenne \bar{Y}_{pi} .

Tant pour l'équation (2) que pour (3) l'estimation des productions totales implique la connaissance des moyennes de troupeau pour les productions partielles (\bar{Y}_p), cependant VAN VLECK et HENDERSON (1961 *f*) ont montré que les estimations par régression intra-population étaient presque aussi efficaces que celles effectuées intra troupeau.

D'après MILLER *et al.* (1972 *a*) une alternative intéressante est proposée par HARVEY en 1959 sous le terme de *régression modifiée*. Ce dernier remarque en effet que

$$bY/Y_p = r_{YY_p} \frac{\sigma Y}{\sigma Y_p}$$

et si les coefficients de variation des productions partielles et totales sont identiques

$$\frac{\sigma Y}{\sigma Y_p} = \frac{\bar{Y}}{\bar{Y}_p}$$

L'équation de régression devient dans ce cas :

$$\widehat{Y} = \bar{Y} + r_{YY_p} \frac{\bar{Y}}{\bar{Y}_p} (Y_p - \bar{Y}_p)$$

soit

$$\widehat{Y} = (1 - r_{YY_p})\bar{Y} + r_{YY_p}CY_p \tag{4}$$

dans laquelle $C = \frac{\bar{Y}}{\bar{Y}_p}$ est le facteur multiplicatif classique.

Cette équation de prévision est donc plus facile à utiliser dans la mesure où seule la moyenne des productions totales du troupeau ou de la population est nécessaire.

Une dernière méthode dite *Méthode P*, proposée par PEARSON et citée par MILLER *et al.* (1972 *b*) constitue également une prévision par régression dont l'originalité consiste à estimer non plus la lactation totale mais la partie inconnue de la lactation par :

$$\widehat{Y}_R = \bar{Y}_R + b(y_p - \bar{y}_p) \tag{5}$$

où Y_R représente le reste de la lactation et y_p une fonction donnée des contrôles connus de la lactation partielle par exemple : $f(1, 2, \dots, m)$ ou $f(2, 3, \dots, m)$ ou bien encore $f(m - 1, m)$. Dans le cas où y_p est l'expression de la production partielle cumulée (Y_p) l'équation (5) correspond à la régression simple, et comme cette dernière la méthode P implique de connaître les moyennes de troupeau pour les productions totales et pour la fonction des résultats partiels. Dans tous les cas la lactation totale est estimée par : $\widehat{Y} = Y_p + \widehat{Y}_R$.

II. — COMPARAISON DES DIFFÉRENTES MÉTHODES

Au cours des 20 dernières années un certain nombre d'auteurs ont étudié l'efficacité comparée de ces différentes méthodes d'extrapolation des lactations partielles en lactations totales et notamment de la méthode multiplicative, la plus utilisée, avec les équations de régression qu'ils proposaient.

I. — *Comparaison des estimations par régression et par facteurs multiplicatifs*

Dans un travail non publié HARVEY (1956) calcule l'espérance théorique de la différence entre les productions totales estimées par des facteurs multiplicatifs d'une part, par régression simple d'autre part, et indique pour cette différence la valeur suivante : $(b - c)(Y_p - \bar{Y}_p)$ où b représente le coefficient de régression de la production totale sur la production partielle Y_p , c la valeur correspondante du facteur multiplicatif et \bar{Y}_p la moyenne des productions partielles cumulées. MADDEN *et al.* (1959) indiquent que la différence $(b - c)$ est négative pendant la quasi totalité de la lactation si bien que l'utilisation de facteurs multiplicatifs a davantage tendance à sous-estimer les mauvaises productions et à surestimer les meilleures. La variance des productions ainsi estimées étant supérieure à celle des estimées par régression d'une valeur en espérance égale à $(b - c)^2 \sigma_{Y_p}^2$. Une telle tendance est effectivement observée par LAMB et MCGILLIARD (1960 a), et SYRSTAD (1964). MADDEN *et al.* (1959) font par ailleurs observer que la variance des estimées par facteurs multiplicatifs est proche de la variance des productions totales réelles et que le classement des animaux étant le même avec toutes les méthodes d'extrapolation il peut être recommandé d'utiliser ces facteurs à des fins de sélection puisqu'ils permettent de maintenir la variance entre groupes d'animaux pour les différents facteurs de variation. Si l'objectif de l'extrapolation des lactations partielles est d'obtenir des estimées des productions réelles aussi précises que possible il convient de comparer les différentes méthodes de ce point de vue. La considération des erreurs moyennes d'une part, de la corrélation entre valeurs estimées et valeurs réelles d'autre part permet de porter un jugement sur la précision des différentes méthodes. Deux études importantes comparent sur un même échantillon les valeurs de ces paramètres pour des estimations de lactations totales par facteurs multiplicatifs et par régression :

Dans la première (APPLEMAN *et al.*, 1969) les coefficients de l'équation de régression sont estimés par moindres carrés avec un modèle faisant intervenir la saison de vêlage et l'intervalle vêlage-saillie fécondante, alors que les estimées par facteurs multiplicatifs n'en tiennent pas compte. La dernière méthode conduit alors à des estimées moyennes légèrement supérieures. La précision, mesurée par R^2 (carré de la corrélation entre valeurs estimées et valeurs réelles) a été calculée par ces auteurs pour des lactations d'ordre 1, 2 et ≥ 3 , et dans chaque cas pour 4 niveaux distincts de production laitière maximum et 8 classes de durée de lactation partielle. Les estimations à partir de 43 000 données portaient à la fois sur les quantités de lait produites et sur les quantités de matière grasse.

Les valeurs de R^2 obtenues avec les estimations par facteurs multiplicatifs étaient inférieures de 0 à 0,08 unités et d'autant plus différentes des estimées par régression que les lactations partielles étaient de courte durée. L'équation de régression exprimait dans ce cas 2,7 p. 100 ; 1,8 p. 100 et 1,2 p. 100 de variance en plus que la part de variation exprimée par l'expression multiplicative, respectivement pour des productions cumulées partielles sur les 2, 3 et 4 premiers mois de lactation. Après le 4^e mois les différences sont très faibles. Les auteurs estiment que ces différences initiales sont probablement liées à la prise en considération de facteurs de variation supplémentaires dans l'équation de régression et que de ce fait les deux méthodes d'extrapolation sont probablement aussi précises. Il faut cependant remarquer que l'estimation de productions totales à l'aide de facteurs multiplicatifs peut être biaisée dans le cas de petits échantillons ou bien si les coefficients de variation des productions partielles et totales dépassent 10 p. 100.

Deux études plus récentes de MILLER *et al.* (1972 a, b) confirment la précision supérieure des méthodes de régression par comparaison avec la méthode multiplicative. L'analyse des erreurs moyennes réelles ($\hat{Y} - Y$) et de leurs valeurs absolues : $|\hat{Y} - Y|$, ainsi que de leur variance sur deux échantillons distincts, l'un extrait du fichier du D. H. I. A. l'autre du troupeau expérimental de l'université de Beltsville, conduit, pour le premier échantillon, aux résultats figurés au tableau I.

TABLEAU I

Moyennes et écart-types des erreurs d'estimation des lactations totales obtenues par application de la méthode multiplicative et la régression modifiée

(D'après MILLER *et al.*, 1972 a)

Nombre de contrôles mensuels connus	Erreur absolue moyenne (kg)		Erreur réelle moyenne (kg)		Écart-type de l'erreur	
	Méthode multiplicative	Régression modifiée	Méthode multiplicative	Régression modifiée	Méthode multiplicative	Régression modifiée
1	1358	793	— 9	— 20	1726	1027
2	772	624	— 27	— 28	990	800
3	571	515	— 23	— 25	731	654
4	471	446	— 18	— 21	601	563
5	405	391	— 9	— 11	513	492
6	348	342	— 5	— 6	437	426
7	294	292	0	— 2	361	356
8	226	228	0	— 1	271	271
9	133	134	— 5	— 4	154	155

Ces chiffres confirment la précision supérieure des estimations par régression surtout pour des lactations partielles de courte durée. Cet avantage était cependant moins manifeste pour les données de Beltsville du fait de corrélations plus élevées entre productions partielles et totales dans cet échantillon, qui en contre partie respectait mieux la condition, nécessaire pour la régression modifiée, d'égalité des coefficients de variation de ces productions.

2. — *Comparaison des méthodes de régression*

Un certain nombre d'auteurs enfin se sont intéressés à la comparaison des différentes méthodes de régression proposées précédemment. Parmi ces études, celles de VAN VLECK et HENDERSON (1961 *b* et *d*) relatives à la comparaison des régressions simples et multiples d'une part, avec ou sans considération de l'effet troupeau d'autre part, sont les plus importantes. Les valeurs obtenues pour les corrélations entre les productions estimées et réelles figurent au tableau 2.

TABLEAU 2

Coefficients de corrélation (R) entre les quantités totales de lait produites et leurs valeurs estimées par régression multiple sur les contrôles connus et par régression simple sur la production partielle cumulée

(D'après VAN VLECK et HENDERSON, 1961 *b*)

Méthode d'extrapolation	Nombre de contrôles mensuels connus								
	1	2	3	4	5	6	7	8	9
Régression multiple	0,57	0,77	0,84	0,88	0,92	0,94	0,96	0,98	0,99
Régression simple	0,57	0,75	0,82	0,87	0,90	0,93	0,95	0,97	0,99

Ces résultats conduisent les auteurs à conclure que la régression simple sur la production cumulée est d'une efficacité si proche de la régression multiple sur chaque contrôle qu'elle est préférable compte tenu de sa plus grande facilité d'emploi. Les variances résiduelles pour des estimations par régression simple intra-troupeau et globale sont présentées au tableau 3.

TABLEAU 3

Variances résiduelles et efficacité comparée d'une extrapolation par régression intra-troupeau (1) et globale (2) sur la production partielle cumulée

(D'après VAN VLECK et HENDERSON, 1961 *f*)

Variances résiduelles	Nombre de contrôles mensuels connus								
	1	2	3	4	5	6	7	8	9
Régression globale (V 2)	4 585	2 829	2 080	1 594	1 176	854	495	325	105
Régression intra-troupeau (V1)	3 774	2 462	1 820	1 407	1 045	763	534	295	96
Efficacité V 2/V 1	121	115	114	113	113	112	111	110	109

La précision supplémentaire de la prévision intra-troupeau est de l'ordre de 20 p. 100 pour des lactations partielles courtes, de 10 p. 100 pour des lactations avec 8 ou 9 contrôles. Ce gain d'efficacité ne justifie cependant pas, selon les auteurs, et dans la plupart des cas, les difficultés de calcul considérables qui résultent de l'application de la méthode.

Dans une étude plus récente de SCHLOTE (1973), qui compare trois variables prédictrices pour la régression : combinaison linéaire des contrôles connus, moyenne de ces contrôles et production journalière moyenne, aboutit également à la conclusion que compte tenu de la précision et de la facilité de calcul, l'emploi d'une régression simple sur la production journalière moyenne sans tenir compte du troupeau constitue une méthode satisfaisante d'extrapolation des lactations partielles. Enfin MILLER *et al.* (1972) dans une étude déjà citée estiment les erreurs absolues moyennes observées par application des méthodes dérivées de la régression simple, c'est-à-dire « régression modifiée » et « méthode P ». Par comparaison à la régression multiple et à la méthode multiplicative, compte tenu de l'information utilisée comme variable prédictrice. Les résultats figurent au tableau 4.

TABLEAU 4

*Erreurs absolues moyennes (kg) pour différentes méthodes d'extrapolation à partir des premiers contrôles mensuels de la lactation partielle (D'après MILLER *et al.*, 1972 b)*

N° d'ordre des contrôles utilisés	Méthode multiplicative	Régression modifiée	Régression multiple	Méthode « P »
1	1 358	862	—	782
1-2	772	641	536	626
2	624	565	—	535
1-3	571	519	431	515
2-3	515	478	—	452
3	538	488	—	429
1-4	471	447	378	—
2-4	449	426	—	403
3-4	445	419	—	385
4	480	449	—	384
1-5	405	392	342	—
2-5	401	387	—	365
3-5	401	383	—	—
4-5	421	403	—	347
5	501	468	—	347

Ils font apparaître l'inefficacité relative de la prévision à l'aide du premier contrôle considéré seul ou avec les contrôles suivants, résultat déjà observé par VAN VLECK et HENDERSON (1961). Pour la plupart des méthodes les contrôles d'ordre 2, 3, 4 et 5 semblent les plus importants du fait de leur corrélation plus élevée avec la production totale. Enfin pour la méthode P le dernier contrôle enregistré semble au moins aussi efficace seul que toute combinaison avec les contrôles précé-

dents, c'est-à-dire qu'il semble être le meilleur indicateur de la production à prévoir, à savoir le reste de la lactation. Cette observation est particulièrement importante dans la mesure où le plus souvent on peut disposer au moment de la prédiction non seulement de la production partielle cumulée mais également du dernier contrôle, ce qui est le cas en France actuellement.

III. — FACTEURS DE VARIATION DES COEFFICIENTS D'EXTRAPOLATION

I. — *Choix de l'échantillon*

Dans la plupart des études citées les coefficients d'extrapolation des lactations partielles ont été estimés à partir de résultats partiels et totaux d'échantillons de lactations complètes de durée égale ou voisine de 305 jours. Or un certain nombre d'animaux sont éliminés précocement pour insuffisance de production en début de lactation. Il est bien évident que la courbe de lactation de ces femelles est différente et que si l'extrapolation de ces lactations partielles est indispensable elle doit être faite avec des facteurs différents de ceux utilisés pour des lactations normales. APPLEMAN *et al.* (1969) souligne l'importance de cette situation pour l'estimation sans biais des taureaux sur descendance puisque, comme nous l'avons signalé, le taux d'élimination précoce de leurs filles est variable selon les taureaux. SYRSTAD (1964) propose, du fait du système norvégien d'expression des performances annuelles de chaque vache, d'estimer un seul jeu de facteurs d'extrapolation à partir d'un échantillon où les animaux sont représentés soit par une lactation partielle soit par une lactation totale. Ainsi les lactations courtes du fait d'une élimination physique de l'animal sont prises en considération au même titre que les lactations normales tronquées à la date du calcul et exprimées en fonction des lactations complètes réalisées la même année.

2. — *Facteurs de variation proprement dits*

De nombreuses études citées par CERSOVSKY (1957) ont mis en évidence l'incidence des facteurs classiques de variation sur la forme de la courbe de lactation donc susceptibles de modifier les coefficients d'extrapolation aussi bien de régression que multiplicatifs.

Age.

La plupart des auteurs déjà cités ont utilisé des lactations d'ordre différent pour tenter de prévoir les productions totales. Tous constatent que l'utilisation de coefficients distincts est nécessaire au moins entre les premières lactations et les suivantes. KEOWN et VAN VLECK (1973) constatent même des coefficients d'extrapolation différents en première lactation pour deux classes d'âge au premier vêlage, confirmant l'observation par AURAN (1973) de l'effet significatif de ce facteur sur la persistance donc sur la relation production partielle-production totale. Selon BAPTIST (1972) et SCHLOTE (1973) cette incidence faible peut être négligée.

Saison de vêlage.

Son effet apparaît significatif dans la plupart des études. McDANIEL *et al.* (1967), BAPTIST (1972), KEOWN et VAN VLECK (1973) estiment ainsi des facteurs d'extrapolation distincts pour chaque saison et SCHLOTE (1973) inclut la saison dans son modèle d'estimation des lactations totales. APPLEMAN (1969) considère quant à lui que cette influence bien que significative n'est pas d'une importance suffisante pour en tenir compte. Seuls FRITZ *et al.* (1960) n'ont pas observé d'effet saison. Dans tous les cas il est inférieur à l'effet du numéro de lactation.

Troupeau, niveau de production.

On observe en moyenne de meilleures persistances des lactations dans les troupeaux de haut niveau, ce qui correspond aussi à des durées de lactation supérieures, POUTOUS et MOCQUOT (1973), SYRSTAD (1964), WOODWARD (1945).

APPLEMAN *et al.* (1969) estiment quant à eux que au niveau individuel, et comme l'avaient fait remarquer HARVEY et KENDRICK, les vaches fortes productrices ont en général un coefficient de persistance inférieur à celui des faibles productrices. Ils constatent d'ailleurs une interaction significative entre le niveau de production maximum et la production aux différents stades de lactation. Dès lors SYRSTAD (1964) propose des facteurs d'extrapolation différents selon le niveau de production de l'étable et APPLEMAN (1969) estime des coefficients distincts en fonction du niveau individuel de production maximum, alors que MADDEN *et al.* (1959), BAPTIST (1972) et SCHLOTE (1973) observent les mêmes relations entre productions partielles et totales pour des fortes et faibles productrices.

Il demeure cependant que l'effet étable de production ne fait pas de doute, même si, comme nous l'avons vu, VAN VLECK et HENDERSON montrent qu'on peut l'ignorer dans une estimation par régression sans réduire considérablement l'efficacité de la prévision. Cette opinion est également exprimée par LAMB et MCGILLIARD (1960), FRITZ *et al.* (1960) ainsi que par BAPTIST (1972).

Race.

Dans les études de LAMB et MCGILLIARD, McDANIEL *et al.*, SCHLOTE, faisant intervenir plusieurs races, les coefficients d'extrapolation des lactations partielles sont différents. Seul BAPTIST considère que cet effet peut être négligé.

Autres facteurs.

Dans la mesure où l'intervalle entre vêlages ou plutôt l'intervalle entre vêlage et saillie fécondante influence la production en fin de lactation, comme l'ont montré de nombreuses études parmi lesquelles celles de SMITH et LEGATES (1962), MILLER et HOOVEN (1969) et AURAN (1974) qui estiment respectivement à 5 p. 100, 2 p. 100 et 7,5 p. 100 la part de variance imputable à ce facteur, les relations entre productions partielles et totales peuvent être modifiées. APPLEMAN *et al.* considèrent cette variable dans l'équation de prévision mais ne constatent pas d'effet significatif. Parmi les autres facteurs susceptibles de modifier la courbe de lactation signalons enfin les différences régionales observées par BAPTIST (1972) et l'influence de l'intervalle vêlage-premier contrôle sur la variance des contrôles successifs estimée à 4,8 p. 100

par AURAN (1973), mais qui n'est pas susceptible de biaiser de façon systématique l'estimation des mâles sur descendance et que l'on peut éliminer en considérant dans l'équation de prédiction non pas le numéro du dernier contrôle de la lactation partielle mais sa durée.

Nous mentionnerons enfin que du fait de la variation du taux butyreux pendant la lactation les facteurs d'extrapolation des productions partielles de matière grasse sont toujours différents de ceux observés pour l'extrapolation des quantités de lait : LAMB et MCGILLIARD (1960 *b*, 1967 *a*, *b*), VAN VLECK et HENDERSON (1961), SYRSTAD (1964), McDANIEL *et al.* (1967), APPELMAN *et al.* (1969), BAPTIST (1972).

CONCLUSIONS

L'intérêt pratique et théorique d'une prévision précoce des productions totales à partir de résultats partiels ne fait aucun doute au moins pour l'indexation rapide plus précise des mâles sur descendance. Une telle extrapolation des lactations partielles est possible compte tenu des liaisons statistiques importantes qui existent entre les productions aux différents stades de la lactation.

Outre les tentatives de formulation algébrique de la courbe de lactation qui conduisent à des équations trop complexes, la plupart des auteurs ont proposé des coefficients simples pour l'extrapolation directe des résultats déjà connus, selon des méthodes qui tentent de rester compatibles avec une réalisation économique des calculs automatiques. La plus simple consiste en l'application de facteurs multiplicatifs estimés préalablement sur un échantillon représentatif. Un certain nombre de coefficients pouvant être estimés pour chaque classe des facteurs de variation les plus importantes. Cette méthode qui n'implique que le stockage des résultats partiels cumulés présente l'avantage de fournir des estimées de même variance que les productions réelles mais ne conduit pas aux estimations les plus précises. Elle risque par ailleurs de favoriser lors de la sélection les productions élevées en début de lactation au détriment de la persistance.

Dans la mesure où on dispose des résultats de tous les contrôles de la lactation partielle une estimation de la production totale par régression multiple semble préférable, bien qu'une régression simple sur la production cumulée soit presque aussi efficace et plus simple à mettre en œuvre.

Les autres méthodes sont dérivées de cette régression simple et conduisent à des estimations plus concordantes avec les productions réelles. Le principal inconvénient de ces méthodes d'estimation par régression réside dans le fait que pour être plus efficace les équations doivent être appliquées intra-troupeau ce qui implique de calculer en même temps les moyennes d'étable pour les productions totales et parfois même pour les productions partielles considérées. Ceci limite considérablement leur possibilité d'utilisation pratique par rapport à la méthode multiplicative qui demeure la plus employée.

On notera enfin avec MILLER *et al.* (1972) que dans la plupart des cas on connaît au moment du calcul la production partielle cumulée à cette date ainsi que la production au dernier contrôle, variable qui semble très liée à la production qu'on cherche

effectivement à estimer, à savoir le reste de la lactation. On peut dès lors se demander si une extrapolation par facteur multiplicatif de la production partielle cumulée, suivie d'une correction additive par régression sur la valeur du dernier contrôle ne constituerait pas une méthode satisfaisante, chaque type de correction corrigeant partiellement les défauts de l'autre, et relativement facile à mettre en œuvre au niveau de la réalisation automatique des calculs.

Il demeure néanmoins évident qu'il faut disposer de coefficients distincts pour les principaux facteurs de variation qui semblent être par ordre décroissant d'importance : numéro de lactation, niveau d'étable, race et saison de vêlage ; aussi bien entendu que pour les différents caractères de production.

Reçu pour publication en décembre 1974.

SUMMARY

STUDIES ON MILK PRODUCTION IN CATTLE.

II. — USEFULNESS OF PART LACTATION RECORDS FOR SELECTION.

a) CRITICAL STUDY OF DIFFERENT METHODS

OF ESTIMATING TOTAL YIELD FROM PART LACTATION RECORDS

Prediction of total milk yield per lactation from partial results has a practical utility for rapid and more precise progeny indexation of males, as well as for early choice of females within-herd. It is possible considering the magnitude of the statistical relations existing between yields at different lactation stages. The various extension methods proposed up to now are described and critically analyzed.

Some authors attempted to formulate a lactation curve equation ; the expressions obtained being too complex, most of the other methods effectively used only employ a limited number of coefficients. Their relative efficacy is compared using two criteria :

- correlation level between estimated and true yields ;
- agreement of variance of estimated and true yields.

The convenience of their utilization for practical calculation is also discussed. The simplest method which is most current is the multiplicative method on cumulative part-yields, the coefficient used being the mean ratio of total yields to part lactation yields of the same duration (same number of monthly tests). It tends to underestimate bad yields and overestimate the best ones. However, it gives estimations of variance close to true variance.

Other methods lead to regression estimates on cumulative part lactation yields (simple regression, modified regression, P method) or on each monthly test (multiple regression). They are more precise in terms of correlation between estimated and true values, but the variance of the estimates is definitely inferior. The simple regression method (or its derivatives) is much easier to use because it only requires keeping one single variable. It is almost as exact as multiple regression. The P method, which attempts to predict the yield remaining to be obtained instead of total yield, is interesting in that the best predicting variable to estimate this yield seems to be last test of the part lactation and it is usually available at the same time.

Calculation of intra herd regression increases the precision of estimates from 10 to 20 p. 100, but it involves the use of herd means for each predicting variable.

Analysis of different variation factors indicates that distinct sets of coefficients must be available, at least for the different numbers of lactations and different characters, and may be also for the various breeds, calving seasons and herd levels.

Finally, extrapolation of « partial » lactations resulting from elimination of the animal during lactation is difficult as the cause of elimination is often unknown. The lactation curve of females culled early for insufficient yield is different and involves the use of coefficients which are appropriate and difficult to estimate.

RÉFÉRENCES BIBLIOGRAPHIQUES

- APPLEMAN R. D., MUSGRAVE D. S., MORRISON R. D., 1969. Extending incomplete lactation records of Holstein cows with varying levels of production. *J. Dairy Sci.*, **52**, 360-368.
- AURAN T., 1973. Studies on monthly and cumulative monthly milk yield records. I. The effect of age, month of calving, herd and length of first test period. *Acta Agr. Scand.*, **23**, 189-199.
- AURAN T., 1974. Studies on monthly and cumulative monthly milk yield records. II. The effect of calving interval and stage in pregnancy. *Acta Agr. Scand.* (sous presse).
- BAPTIST R., 1972. Zur Hochrechnung von Teillaktationen bei Färsen. *Züchtungskunde*, **44**, 295-306.
- CANNON C. Y., FRYE J. B. jr, SIMS J. A., 1942. Predicting 305 days yield from short-time records. *J. Dairy Sci.*, **35**, 991-999.
- CERSOVSKY H., 1957. Der wirtschaftliche und züchterische Wert des Verlaufes der Laktationskurve bei Milchkühen. *Tierzucht*, **11**, 185-188.
- FRITZ G. R., MCGILLIARD L. D., MADDEN D. E., 1960. Environmental influence on regression factors for estimating 305-days production from part lactations. *J. Dairy Sci.*, **43**, 1108-1117.
- GAINES W. L., 1927. Persistency of lactation in dairy cows. *Ill. agric. Exp. Stn.*, Bull. n° 288.
- GUSTAFSON G., 1972. Appropriate part of the lactation for progeny testing in milk production with special regard to the current pregnancy. *Lantbr Högsk. Meddn.*, Serie A, n° 168.
- HARVEY W. R., 1956. Extension of incomplete records to a 10 month basis (données non publiées).
- HARVEY W. R., 1959. Problems to consider in determining appropriate extension factors for incomplete records (données non publiées).
- HOFFMANN G., 1959. Über die Brauchbarkeit von Färsenkurzleistungen für die Erbwertermittlung. *Züchtungskunde*, **31**, 210-217.
- KEOWN J. F., VAN VLECK L. D., 1971. Selection on test-day fat percentage and milk production. *J. Dairy Sci.*, **54**, 199-203.
- KEOWN J. F., VAN VLECK L. D., 1973. Extending lactation records in progress to 305-days equivalent. *J. Dairy Sci.*, **56**, 1070-1079.
- LAMB R. C., MCGILLIARD L. D., 1959. Variables affecting the ratio of total milk produced in 305 days to monthly milk production. *J. Dairy Sci.*, **42**, 924.
- LAMB R. C., MCGILLIARD L. D., 1960 a. Variables affecting ratio factors for estimating 305-days production from part lactations. *J. Dairy Sci.*, **43**, 519-528.
- LAMB R. C., MCGILLIARD L. D., 1960 b. Comparison of ratio factors for extending part-time milk and fat records. *J. Dairy Sci.*, **43**, 879.
- LAMB R. C., MCGILLIARD L. D., 1967 a. Ratio factors to estimate 305-days production from lactation records in progress. *J. Dairy Sci.*, **50**, 1101-1108.
- LAMB R. C., MCGILLIARD L. D., 1967 b. Usefulness of part records to estimate the breeding values of dairy cattle. *J. Dairy Sci.*, **50**, 1458-1467.
- MADDEN D. E., LUSH J. L., MCGILLIARD L. D., 1955. Relations between parts of lactations and production ability of Holstein cows. *J. Dairy Sci.*, **38**, 1264-1271.
- MADDEN D. E., MCGILLIARD L. D., RALSTON N. P., 1956. Relations between monthly test-day milk production of Holstein-Friesian cows. *J. Dairy Sci.*, **39**, 932.
- MADDEN D. E., MCGILLIARD L. D., RALSTON N. P., 1959. Relations between test-day milk production of Holstein cows. *J. Dairy Sci.*, **42**, 319-326.
- MCDANIEL B. T., MILLER R. H., CORLEY E. L., 1967. Sources of variation in ratios of total to part yield. *J. Dairy Sci.*, **50**, 1917-1924.
- MILLER R. H., MCDANIEL B. T., CORLEY E. L., 1967. Variation in ratio factors for age-adjusting part-lactation records. *J. Dairy Sci.*, **50**, 1819-1823.
- MILLER R. H., HOOVEN N. W. jr, 1969. Factors affecting whole-and part lactation milk yield and fat percentage in a herd of Holstein cattle. *J. Dairy Sci.*, **52**, 1588-1600.
- MILLER R. H., HOOVEN N. W. jr, SMITH J. W., HARVEY W. R., CREEGAN M. E., 1972 a. Modified regression for estimating total lactation from part lactation yields. *J. Dairy Sci.*, **55**, 208-213.
- MILLER R. H., PEARSON R. E., FOHRMAN M. H., FREEGAN M. E., 1972 b. Methods of projecting complete lactation production from part-lactation yield. *J. Dairy Sci.*, **55**, 1602-1606.
- M. M. B. (Milk Marketing Board) 1971-1972. *The practical application of the lactation curve equation*.
- MOCQUOT J. C., POUTOUS M., 1971. Les relations entre les index laitiers des pères et des fils des races bovines. *F. F. P. N., Montbéliarde et Normande. Ann. Génét. Sél. anim.*, **3**, 201-206.
- MOXLEY J. E., 1968. Within herd evaluation in the dairy herd analysis service program. (données non publiées).
- NELDER J. A., 1966. Inverse polynomials, a useful group of multi-factor response functions. *Biometrics*, **22**, 128-141.

- POUTOUS M., MOCQUOT J. C., 1973. Études sur la production laitière des bovins. I. Note sur la correction du niveau d'étable. *Ann. Génét. Sél. anim.*, **5**, 211-216.
- SCHLOTE W., 1973. Die Aufrechnung und Verwendung unvollständiger Färsen laktationen für die Zuchtwertschätzung von Bullen. *Z. für V. Zuchtbiol.*, **89**, 265-279.
- SEARLE S. R., 1961 b. Part lactations. II. Genetic and phenotypic studies of monthly milk fat yield. *J. Dairy Sci.*, **44**, 282-295.
- SMITH J. W., LEGATES J. E., 1962. Relation of days open and days dry to lactation milk and fat yield. *J. Dairy Sci.*, **45**, 1192-1198.
- SYRSTAD O., 1964. Studies on dairy herd records. I. Evaluation of incomplete records. *Acta Agr. Scand.*, **14**, 129-149.
- VAN VLECK L. D., 1962. Effect of incomplete records on sire evaluation. *J. Dairy Sci.*, **45**, 1511-1515.
- VAN VLECK L. D., HENDERSON C. R., 1961 a. Estimates of genetic parameters of some functions of part lactation milk records. *J. Dairy Sci.*, **44**, 1073-1084.
- VAN VLECK L. D., HENDERSON C. R., 1961 b. Regression factors for extending part lactation milk records. *J. Dairy Sci.*, **44**, 1085-1092.
- VAN VLECK L. D., HENDERSON C. R., 1961 c. Ratio factors for adjusting monthly test-day data for age and season of calving and ratio factors for extending part lactation records. *J. Dairy Sci.*, **44**, 1093-1102.
- VAN VLECK L. D., HENDERSON C. R., 1961 d. Regression factors for predicting a succeeding complete lactation milk record from part lactation records. *J. Dairy Sci.*, **44**, 1322-1327.
- VAN VLECK L. D., HENDERSON C. R., 1961 e. Use of part lactation records in sire evaluation. *J. Dairy Sci.*, **44**, 1511-1518.
- VAN VLECK L. D., HENDERSON C. R., 1961 f. Extending part lactation records by regression ignoring herd effects. *J. Dairy Sci.*, **44**, 1519-1528.
- VAN VLECK L. D., HENDERSON C. R., 1961 g. Utilizing both part and complete daughter records in sire evaluation. *J. Dairy Sci.*, **44**, 2068-2076.
- VUJICIC I., BACIC B., 1961. Nouvelle equation de la courbe de lactation (Croatie). *Novi Sad Ann. Sci. Agric.*, n° 5.
- WOOD P. D. P., 1967. Algebraic model of the lactation curve in cattle. *Nature*, London, **216**, 164-165.
- WOOD P. D. P., 1969. Factors affecting the shape of the lactation curve in cattle. *Anim. Prod.* **11**, 307-316.
- WOODWARD T. E., 1945. Some studies on lactation records. *J. Dairy Sci.*, **28**, 209-218.
-