



**HAL**  
open science

## Description de la base de données "LASCAR", un outil pour l'étude des carrières des vaches laitières

J-B. Coulon, Françoise Lescourret, B. Faye, E. Landais, J-L. Troccon, L. Pérochon

### ► To cite this version:

J-B. Coulon, Françoise Lescourret, B. Faye, E. Landais, J-L. Troccon, et al.. Description de la base de données "LASCAR", un outil pour l'étude des carrières des vaches laitières. *Productions Animales*, 1993, 6 (2), pp.151-160. hal-00896051

**HAL Id: hal-00896051**

**<https://hal.science/hal-00896051v1>**

Submitted on 11 May 2020

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

J-B. COULON, Françoise LESCOURRET\*,  
B. FAYE\*, E. LANDAIS\*\*,  
J-L. TROCCHON\*\*\*, L. PÉROCHON\*

INRA Laboratoire Adaptation des  
Herbivores aux Milieux  
Theix, 63122 St-Genès-Champanelle

\* INRA Laboratoire d'Ecopathologie  
Theix, 63122 St-Genès-Champanelle

\*\* INRA Unité SAD Versailles-Dijon-  
Mirecourt

Route de St-Cyr 78026 Versailles Cedex

\*\*\* INRA Station de Recherches sur la  
Vache Laitière 35590 St-Gilles

# Description de la base de données "LASCAR", un outil pour l'étude des carrières des vaches laitières

Si de très nombreux travaux expérimentaux ont été consacrés aux facteurs de variation de la production laitière à court terme (1 à 3 mois) ou même à l'échelle d'une lactation complète (Broster et Broster 1984), les études entreprises sur de longues périodes sont beaucoup plus rares (Wiktorsson 1979, Coulon *et al* 1990). Pourtant, elles ont permis de mettre en évidence des effets cumulatifs, dégressifs ou compensateurs des facteurs de variation (en particulier alimentaires) qui peuvent modifier sensiblement les conclusions tirées des observations réalisées sur des périodes de courte durée. Par ailleurs, le cloisonnement des disciplines et/ou la durée limitée des périodes expérimentales conduisent le plus souvent à négliger, dans l'étude des performances de production, l'effet de phénomènes passés ou concomitants exprimés par des performances d'un autre ordre. Ainsi, les résultats de reproduction et de pathologie sont-ils rarement intégrés parmi les facteurs potentiellement explicatifs de la production laitière. Réciproquement, une partie des études d'épidémiologie vétérinaire ne prend pas en compte les caractères de productivité des animaux (Dohoo *et al* 1982-1983, Enevoldsen *et al* 1991). Pourtant, quelques travaux ont démontré l'effet notable des facteurs sanitaires sur la production laitière (Lucey *et al* 1986) et leurs interrelations avec la reproduction (Coulon *et al* 1989) ou à l'inverse les conséquences pathologiques des

hauts niveaux de production (Barnouin et Karaman 1986). A ce titre, les études d'écopathologie, qui prennent en compte à la fois les paramètres de l'environnement biologique, physique, voire humain et économique, susceptibles d'induire des troubles sanitaires chez les animaux d'élevage, permettent d'étudier les relations entre les différents types de performances (Ganière *et al* 1991).

L'étude conjointe des différentes performances de la vache laitière sur le long terme nous paraît donc potentiellement fructueuse. **La carrière d'un animal domestique** peut être définie comme l'enchaînement des événements individuels qui permettent de caractériser sa vie sur le plan de la croissance, de la production, de la reproduction et de la santé, de sa naissance à sa mort ou à sa réforme (Lasseur et Landais 1992). Elle peut être vue comme la résultante d'interactions complexes entre les caractéristiques propres de l'animal (essentiellement génétiques) et celles relatives à son environnement (milieu d'élevage et mode d'élevage). Un travail engagé dans ce sens sur la vache laitière a mis en évidence la variabilité individuelle des performances à l'échelle de la lactation et de la carrière et suggéré l'existence de types de carrières particuliers liés entre autres à la race des animaux et à leur alimentation (Coulon *et al* 1990). Ces différentes hypothèses doivent cependant être vérifiées et précisées sur une gamme plus large de situations et avec des effectifs plus élevés. Un programme de recherche a donc été élaboré dans ce sens pour décrire correctement le déroulement de la carrière d'une vache laitière et ses principaux facteurs de variation, et mettre en évidence des indicateurs précoces de sa durée et de ses caractéristiques. Des retombées intéressantes peuvent être attendues en termes de recommandations techniques pour la conduite des troupeaux et la gestion des carrières des animaux (politique de réforme et de renouvellement, choix du matériel génétique).

Le centre INRA de Clermont-Fd-Theix offre une opportunité exceptionnelle de travail sur les carrières des animaux. Il existe en effet un nombre considérable de données

## Résumé

Dans l'objectif d'étudier la carrière des vaches laitières (à partir du premier vêlage), une base de données relationnelle, baptisée LASCAR, a été élaborée. Elle a été alimentée à partir d'enregistrements réalisés au cours des 20 dernières années dans les troupeaux laitiers des 3 domaines expérimentaux du centre INRA de Clermont-Fd-Theix. Elle est composée de 10 tables décrivant les caractéristiques génétiques et de la période d'élevage des animaux, les performances individuelles à différentes échelles de temps (production et composition du lait, poids vif, troubles sanitaires, événements liés à la reproduction), et quelques facteurs de variation de ces performances (conduite alimentaire, climat). Elle regroupe 3851 lactations provenant de 1179 carrières. Ses données ont fait l'objet d'un contrôle de cohérence exhaustif. Différentes démarches d'étude de la carrière à partir de cette base sont présentées.

### Définition et intérêts d'une base de données

Les bases de données sont des structures d'archivage informatisées permettant un accès souple et rapide à l'information (Williams 1986, Udomprasert et Williamson 1990). Leurs fonctions les rendent particulièrement intéressantes dans des programmes de recherche exploratoire multi-objectifs manipulant des données nombreuses et de nature diverse.

La réalisation technique d'une base de données doit être précédée de trois étapes de conception : élaboration d'un "dictionnaire de données" (liste des données nécessaires), modélisation conceptuelle des données (représentation sémantique de l'information requise, à un niveau purement abstrait, qui ne préjuge donc pas de choix techniques), conception technique (traduction du modèle conceptuel tenant compte de choix techniques, par exemple construction d'un schéma relationnel) (Rolland 1986). Ces étapes peuvent être menées à bien grâce à des méthodes de conception de système d'information telles que MERISE (Tardieu *et al* 1983). La traduction d'un modèle conceptuel en schéma de base de données est pilotée par un ensemble de règles qui la rendent triviale. Le schéma relationnel est un ensemble de tables qui sont des collections de types de données élémentaires (par exemple : numéro d'identification, race, date de réforme ...). Les liens entre tables sont établis lors de l'interrogation de la base à partir de leurs données communes. Enfin, dans des bases de données servant des projets de recherche exploratoire, la connaissance de grands principes relatifs aux futurs traitements des données peut influencer utilement le modèle conceptuel et sa traduction en schéma de base de données (Lescourret *et al* 1992).

L'attention apportée à la conception, qui nécessite un dialogue constant avec les futurs utilisateurs de la base, garantit **une bonne maîtrise du système d'informations** (données, besoins) qui prévient de coûteuses erreurs (volumes utilisés, données non introduites ...) (Hall 1978). Le modèle conceptuel, qui peut être présenté sous forme graphique, est une image fidèle de l'état du système, utile en phase d'exploitation de la base. Par ailleurs, une base de données garantit une **forte sécurité d'utilisation** : les règles que doivent suivre les données pour être cohérentes avec la réalité et les objectifs fixés sont ou bien contenues dans le schéma de la base, ou bien facilement vérifiables avec les Langages de Manipulation de Données (LMD) des Systèmes de Gestion de Bases de Données (SGBD). Un point crucial est **la facilité et la rapidité de l'accès à l'information** qui dépendent au moins autant de la pertinence de la structure mise en place que de la puissance des LMD. Enfin, l'analyse exploratoire de données nécessite des tableaux de données nombreux et répondant à des critères de sélection variés. Il faut par ailleurs revenir en permanence aux données. Cette démarche est grandement facilitée par l'utilisation d'une base de données relationnelle. Celle-ci garantit une bonne indépendance entre données et traitements, au contraire d'un ensemble de fichiers indépendants.

concernant à la fois les caractéristiques originelles des animaux, leur croissance dans le jeune âge, leurs performances de production, leur santé et leur reproduction au cours de leurs lactations successives, ainsi que les causes de leur réforme. Sur les troupeaux laitiers des 3 domaines expérimentaux dépendant du Centre (Theix, Orcival et Marcenat), ces données ont pour la plupart été relevées systématiquement et saisies en grande partie sur support informatique selon un cadre identique depuis pratiquement 20 ans. Pour faciliter la réalisation des objectifs poursuivis, il a été décidé de les organiser au préalable dans une base de données. La définition et les intérêts de ce type de structuration sont précisés en encadré. Cette base a été baptisée LASCAR (Lait, Santé, Carrières, Reproduction). Elle est un exemple de structuration de données adaptée à un projet scientifique. C'est une base non évolutive fortement orientée vers le traitement des données.

L'objectif de cet article est de présenter la structure et le contenu de cette base et d'évoquer les grandes lignes de son utilisation.

### Choix et définition des variables

Les variables retenues décrivent :

- les performances zootechniques (production et composition du lait, poids vif), mesurées

- à l'échelle de la lactation (taux butyreux et protéique) et/ou de la semaine de lactation (lait, poids vif) ;

- certaines caractéristiques individuelles de l'alimentation. En l'absence, pour les 2/3 des lactations concernées, de données informatisées concernant les quantités de fourrages ingérées et les bilans nutritifs (qui auraient été d'un intérêt certain), seules les quantités de concentrés ingérées à l'échelle hebdomadaire ont été retenues. Le niveau des apports énergétiques et azotés en début de lactation (semaines 1 à 8 et 9 à 16) a été estimé individuellement à partir des protocoles concernant les expériences pour lesquelles les vaches avaient été utilisées. Une échelle à trois niveaux (bas, moyen et haut, par rapport aux recommandations de l'INRA) a été jugée suffisante. Le niveau moyen a été affecté aux vaches non expérimentales ;

- les événements subis par les vaches durant leur carrière: troubles sanitaires, pratiques prophylactiques individuelles, événements liés à la reproduction (inséminations, vélages, caractéristiques du veau), tarissements, réforme ;

- des attributs invariables à l'échelle de la carrière, et déterminés à son début: race, descripteurs de la parenté et indicateur de la valeur génétique (index laitier du père

Tableau 1. Nombres de données de la base relatives à quelques variables majeures, ventilés par domaine.

Domaine	Marcenat	Theix	Orcival	Total
Période (1)	1/10/67 1/6/88	1/10/67 13/9/87	1/10/79 30/9/91	
Vaches	538	335	306	1179
dont				
Montbéliardes	267	1	0	268
Française-Frisonnes	209	90	0	299
Holstein	13	162	158	333
FF x Holstein	45	71	53	169
autres <sup>(2)</sup>	4	11	95	110
Lactations	1805	1173	873	3851
Avortements <sup>(3)</sup>	46	30	11	87
Troubles sanitaires <sup>(4)</sup>	2161	2737	1092	5990
dont				
mammites	497	814	379	1690
panaris	756	139	43	938
fièvre vitulaire	36	110	58	204
Mesures de poids vif à un an	505	301	274	1080
Nb de vaches dont on connait l'index du père	385	249	222	856
Mesures de climat <sup>(5)</sup>	18	17	6	41

(1) domaine d'occurrence des dates de vêlage.

(2) autres : principalement Montbéliarde x Holstein

(3) inducteur d'une lactation ou non.

(4) non comprises les interventions prophylactiques, qui font aussi partie des événements sanitaires.

(5) nombre de campagnes bien renseignées.

(Institut de l'Élevage-INRA 1992a et b)), et certaines caractéristiques de la période d'élevage des animaux (poids, inséminations...);

- le climat des domaines expérimentaux, sous la forme de quelques descripteurs annuels, élaborés à partir des données journalières de la banque de données climatiques INRA.

Par ailleurs, parmi les variables décrivant les conduites d'élevage (modalité du pâturage, type de stabulation ...), seules les bornes des périodes de pâturage et de récolte des fourrages ont été retenues pour figurer dans la base. Les autres variables ont été décrites dans un document d'accompagnement retraçant les principaux événements qui ont marqué la vie des domaines pendant la période d'étude (Pérochon 1990).

La production laitière, les poids vifs, les troubles sanitaires, les événements liés à la reproduction et les caractéristiques de la réforme constituent les "performances" étudiées. Les attributs invariables à l'échelle de la carrière et déterminés à son début, les pratiques (conduites d'élevage, pratiques alimentaires et sanitaires) et le climat constituent les facteurs de variation de ces performances.

Le dictionnaire des données et la liste des contraintes suivies par ces données (par exemple: une vache a appartenu à un seul domaine durant sa carrière) ont été établis simultanément. Les périodes d'étude et les caractéristiques des animaux retenus ont été choisies en fonction des disponibilités offertes par chacun des domaines. Les périodes d'étude couvrent les campagnes 1968-1988 pour les domaines de Marcenat et de Theix, et les campagnes 1980-1991 pour le domaine d'Orcival. Ont été retenues les vaches ayant réalisé dans un même domaine une succession ininterrompue de lactations de la première mise bas jusqu'à la réforme, ainsi que 89 carrières tronquées par la fin de la période d'étude. Les vaches fistulées (domaine de Theix) ont été exclues. 1179 vaches ayant réalisé 3851 lactations ont ainsi été retenues (tableau 1).

## Organisation de LASCAR

Le modèle conceptuel de LASCAR (figure 1) a été décrit par ailleurs (Lescourret *et al* 1992). Il a été traduit en un ensemble de 10 tables, présentées au tableau 2.

Les tables VACHE, LACTATION et SEM-LACT (respectivement 1179, 3851 et 182678

enregistrements) réunissent les mesures uniques, à l'échelle de la carrière, de la lactation et de la semaine de lactation respectivement. Les tables INS (5122 enregistrements) et SUBIT (6129 enregistrements) décrivent les événements relatifs aux inséminations et aux événements sanitaires respectivement. La table TYPEVEN (81 enregistrements) décrit le regroupement des événements sanitaires observés en grands types (par exemple, les événements "boiterie", "cerise", "amputation d'onglon" et "limace" sont regroupés dans le type "boiterie"). La table VEAU (3922 enregistrements) contient les caractéristiques des veaux nés des vaches de VACHE. La table

TAUREAU (147 enregistrements) contient les numéros et les index laitiers des pères des vaches de VACHE, s'ils sont indexés (base 1992). La table CLIMPATUR (54 enregistrements) comprend, pour chaque domaine, les bornes des périodes de pâturage pour chacune des campagnes incluses dans sa période d'étude, et la liste des paramètres climatiques caractérisant ces campagnes. Enfin, la table DOMAINE (2 enregistrements) contient les dates moyennes de début et de fin des périodes de récolte du foin et de l'ensilage. Au total la base contient environ 1,5 millions de données.

La structure des tables est adaptée à des besoins spécifiques en traitements de données.

Figure 1. Modèle conceptuel de LASCAR : organisation simplifiée des types de données en entités (carrés) et relations (cercles). Les entités "vache" et "date" sont représentées 2 fois. Les couples de nombres associés aux relations représentent les nombres minimum et maximum d'implications des entités dans les relations ; par exemple une cause de réforme donnée peut s'appliquer à une vache au minimum et à N vaches au maximum (N représente tout nombre supérieur à 2).

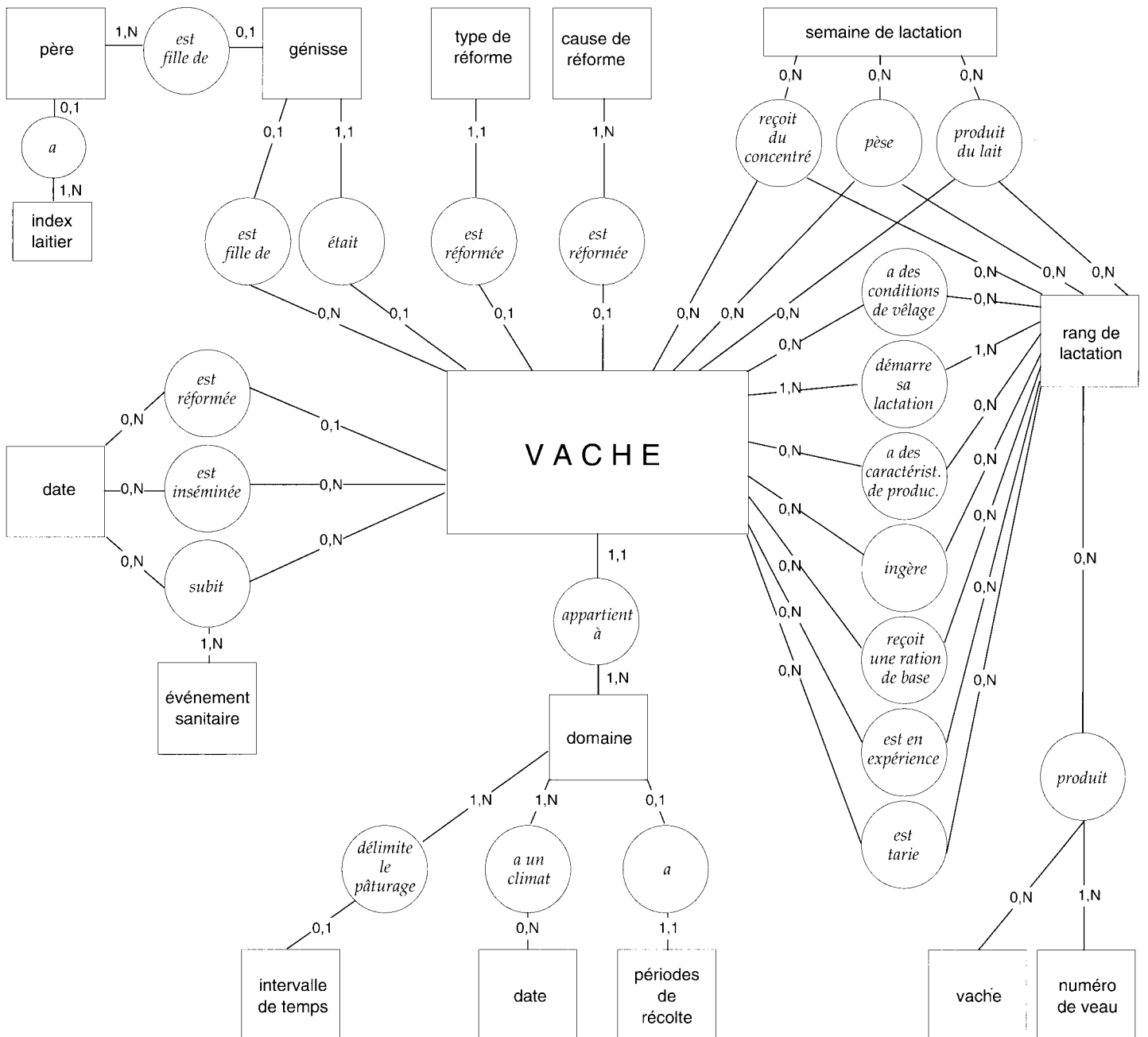


Tableau 2. Tables de LASCAR (dans chaque table, les clés primaires sont en italiques).

<p><b>VACHE</b></p> <p><i>numéro</i></p> <p>date de naissance</p> <p>race</p> <p>type de réforme</p> <p>cause de réforme</p> <p>date de sortie</p> <p>rang de lactation de sortie</p> <p>semaine de lactation de sortie</p> <p>domaine</p> <p>rang de lactation de la mère</p> <p>difficulté de naissance</p> <p>naissance gémellaire</p> <p>date de sevrage</p> <p>date de première insémination</p> <p>date de première insémination fécondante</p> <p>numéro de cette insémination</p> <p>poids vif à la naissance</p> <p>poids vif au sevrage</p> <p>poids vif à 180 jours</p> <p>poids vif à 360 jours</p> <p>poids vif à la première insémination fécondante</p> <p>numéro de la mère</p> <p>numéro du père</p> <p><b>INS</b></p> <p><i>numéro de vache</i></p> <p><i>date</i></p> <p>numéro de lactation</p> <p>semaine de lactation</p> <p>numéro de l'insémination</p> <p>chaleurs ou synchronisation</p> <p>saillie ou insémination artificielle</p> <p>résultat de l'insémination</p> <p><b>SUBIT</b></p> <p><i>numéro de vache</i></p> <p><i>date</i></p> <p><i>événement</i></p> <p>numéro de lactation</p> <p>semaine de lactation</p> <p>indice de gravité</p> <p><b>TYPEVEN</b></p> <p><i>événement</i></p> <p>catégorie de cet événement</p> <p><b>VEAU</b></p> <p><i>numéro de vache</i></p> <p><i>numéro de lactation</i></p> <p><i>numéro d'ordre</i></p> <p>sexe</p> <p>survie</p> <p>poids vif</p> <p><b>DOMAINE</b></p> <p><i>nom du domaine</i></p> <p>date moyenne du début de récolte du foin</p> <p>date moyenne de la fin de récolte du foin</p> <p>date moyenne du début de récolte de l'ensilage</p> <p>date moyenne de la fin de récolte de l'ensilage</p>	<p><b>LACTATION</b></p> <p><i>numéro de vache</i></p> <p><i>numéro de lactation</i></p> <p>date de début de lactation</p> <p>numéro de la dernière semaine de la lactation antérieure</p> <p>type d'induction (avortement ou vêlage)</p> <p>état au vêlage</p> <p>heure de vêlage</p> <p>difficulté de vêlage</p> <p>présence de jumeaux</p> <p>production des 4<sup>e</sup>, 5<sup>e</sup> et 6<sup>e</sup> jours de lactation</p> <p>apport énergétique semaines 1-8</p> <p>apport azoté semaines 1-8</p> <p>apport énergétique semaines 9-16</p> <p>apport azoté semaines 9-16</p> <p>taux butyreux moyen</p> <p>taux protéique moyen</p> <p>production laitière totale</p> <p>production laitière en 305 jours</p> <p>durée de production</p> <p>persistance hivernale</p> <p>persistance estivale</p> <p>production laitière semaines 1-8</p> <p>production laitière semaines 1-16</p> <p>type de ration de base hivernale</p> <p>nombre d'expériences</p> <p>date de tarissement</p> <p>semaine de tarissement</p> <p><b>SEMLACT</b></p> <p><i>numéro de vache</i></p> <p><i>numéro de lactation</i></p> <p><i>semaine de lactation</i></p> <p>date du milieu de cette semaine</p> <p>poids vif hebdomadaire</p> <p>production laitière hebdomadaire</p> <p>concentré hebdomadaire</p> <p><b>TAUREAU</b></p> <p><i>numéro</i></p> <p>index laitier</p> <p><b>CLIMPATUR</b></p> <p><i>nom du domaine</i></p> <p><i>date de mise à l'herbe</i></p> <p><i>date de rentrée à l'étable</i></p> <p>somme des températures positives (20/2 au 20/6)</p> <p>somme des ETR décadaires (21/6 au 10/10)</p> <p>nombre de séquences sèches durant la fenaison</p> <p>nombre de jours avec pluie &gt; 10 mm au printemps</p> <p>nombre de jours avec sol couvert de neige au printemps</p> <p>nombre de jours avec pluie &gt; 10 mm en été</p> <p>nombre de jours avec sol couvert de neige en été</p> <p>type de climat printanier (1)</p> <p>type de climat estival (1)</p> <p>type de climat de novembre à janvier (1)</p> <p>type de climat de février à avril (1)</p>
---	---

(1) élaboré à partir des données mensuelles (température, pluviométrie ...) fournies par le STEFCE.

Ainsi, le besoin de disposer à la fois du temps calendaire et du temps physiologique (rang et semaines de lactation) pour les événements et les mesures subis par les vaches a conduit à introduire les deux types de temps dans la plupart des tables. De même, des productions laitières totales et des coefficients de persistance de la lactation, susceptibles d'être fréquemment requis, ont été calculés à partir des données de la table SEMLACT, et introduits dans la table LACTATION. Dans ces deux cas, on a amendé la traduction brute du modèle conceptuel en introduisant dans quelques tables des redondances nécessaires (données ne pouvant être obtenues par ailleurs que par un calcul fastidieux à partir de la base).

Les liens entre tables sont établis lors de l'interrogation, par des "jointures" basées sur leurs champs communs. Ainsi, les tables LACTATION, SUBIT, INS, SEMLACT et VEAU possèdent en commun les champs "numéro de vache" et "numéro de lactation" qui permettent d'associer, dans un tableau de sortie portant en ligne des lactations, des caractéristiques de production, de santé et de reproduction.

Cette base, installée sur le serveur du Centre de Theix, est gérée par le SGBD Relationnel ORACLE. Elle peut être interrogée à l'aide du langage SQL (Structured Query Language), le langage standard d'interrogation des bases de données relationnelles, qui est assez proche du langage naturel. Compte tenu des configurations locales, de la simplicité du modèle et de l'indexation des tables, les temps d'exécution sont courts malgré des nombres d'enregistrements parfois importants. Par exemple, l'obtention de la liste des numéros de vache associés à une variable prenant la valeur 1 en cas de survenue d'un au moins des troubles sanitaires parmi une liste de 6 pendant la première lactation, et 0 sinon, demande seulement 25 secondes sur le serveur du centre de Theix. L'obtention d'une série de numéros de vache associés à des caractéristiques initiales de ces animaux (race, campagne de naissance, poids à un an, index laitier du père), de leur première lactation (âge au premier vêlage, production initiale, quantité moyenne de lait produite par jour), et de leur carrière (durée, rang de lactation de sortie, type et cause de réforme), nécessite seulement 60 secondes. Cette requête met en jeu les tables VACHE, LACTATION et TAUREAU et impose des conditions de tri relatives au numéro de lactation, à l'existence d'un index laitier renseigné et au fait que la carrière soit complète. Elle fait appel à 14 des variables de ces tables.

### Validation de LASCAR

Pour augmenter la fiabilité d'une base, on doit s'assurer que les données qu'elle contient respectent des règles de cohérence. Ces règles peuvent concerner :

1) une seule variable (domaine d'occurrence). Par exemple, le poids vif des animaux dans la table SEMLACT doit être compris entre 400 et 900 kg.

2) plusieurs variables d'une même table. Dans la table LACTATION, la variable "production des 4<sup>e</sup>, 5<sup>e</sup> et 6<sup>e</sup> jours de lactation" ne doit pas être renseignée si la lactation dure moins de 6 jours.

3) plusieurs variables dans des tables différentes. Par exemple la variable "type d'induction" de la table LACTATION ne peut pas prendre la valeur 1 si la durée de gestation calculée à partir de la date d'insémination présumée fécondante (variables "date" et "résultat" de la table INS) et de la date de vêlage (variable "date de début" de la table LACTATION) est inférieure à 240 jours ; il s'agit alors d'un avortement, et la variable "type d'induction" prend la valeur 0. Les règles de cohérence concernent aussi les données "élaborées" de la base (c'est à dire construites à partir de données brutes de la base). Ainsi, la variable "durée de production" de la table LACTATION doit être égale à la différence entre les variables "date de début de lactation" et "date de tarissement" de la même table, ou, en l'absence de tarissement, entre les 2 variables "date de début de lactation" relatives à 2 rangs de lactation successifs, ou, pour la dernière lactation et en l'absence de tarissement, à la différence entre les variables "date de début de lactation" de la table LACTATION et "date de sortie" de la table VACHE.

La vérification de ces règles a été assurée à différentes étapes de l'élaboration de la base. Dans un premier temps, les fichiers originels ont fait l'objet de vérifications manuelles et/ou informatiques, suivies de corrections éventuelles. Les valeurs aberrantes des taux protéique et butyreux annuels ont ainsi été détectées. Dans un second temps, après l'alimentation de la base, 127 requêtes en langage SQL ont été lancées pour vérifier la cohérence des données. Les erreurs détectées ont été corrigées soit immédiatement (par exemple élimination des doublons) ou après retour aux données de base, voire expertise des utilisateurs. Cette dernière phase a nécessité environ 3 semaines de travail à 2 personnes, ce qui n'est pas considérable compte tenu de la quantité de données et de la sécurité garantie pour l'utilisation ultérieure.

### Utilisation de LASCAR

La répartition des principaux types de données par domaine est précisée au tableau 1, et la plage de variation de certaines d'entre elles à la figure 2. Cette plage apparaît très étendue. Ainsi, 611 vaches ont réalisé au moins 3 lactations (soit 56 %) et 176 au moins 6 (soit 16 %). Au cours de ces lactations, la production laitière totale a varié de 0 à 14528 kg (5032 kg en moyenne). 76 % de ces productions ont été comprises entre 3000 et 7000 kg. Les poids au premier vêlage ont été compris entre 400 et 700 kg.

Compte tenu de ces caractéristiques, LASCAR nous paraît pouvoir répondre à l'objectif d'étude des carrières des vaches laitières. Cependant, une difficulté majeure réside dans le choix d'échelles de temps différentes pour

Figure 2. Amplitude de variation de quelques types de données de LASCAR.

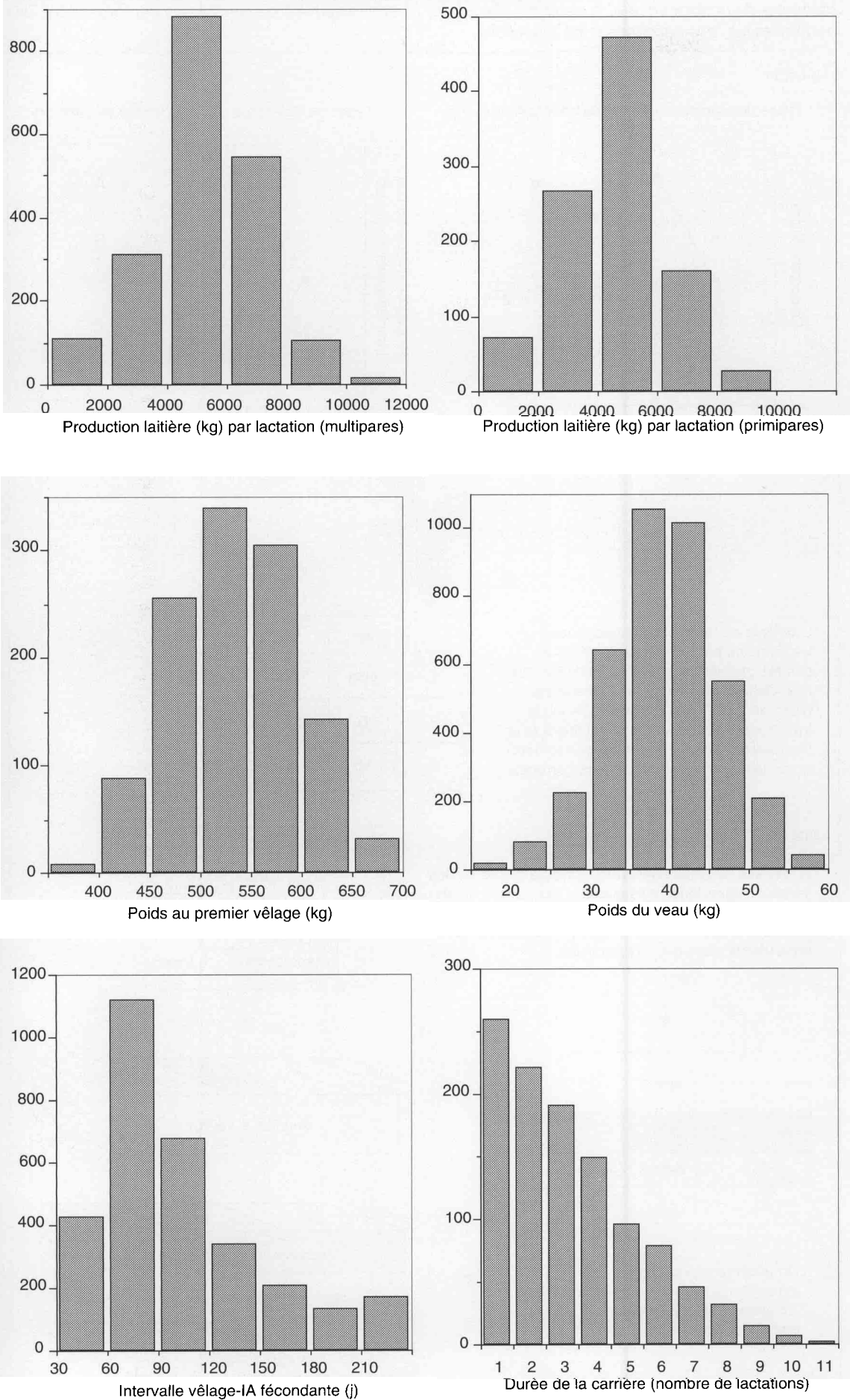
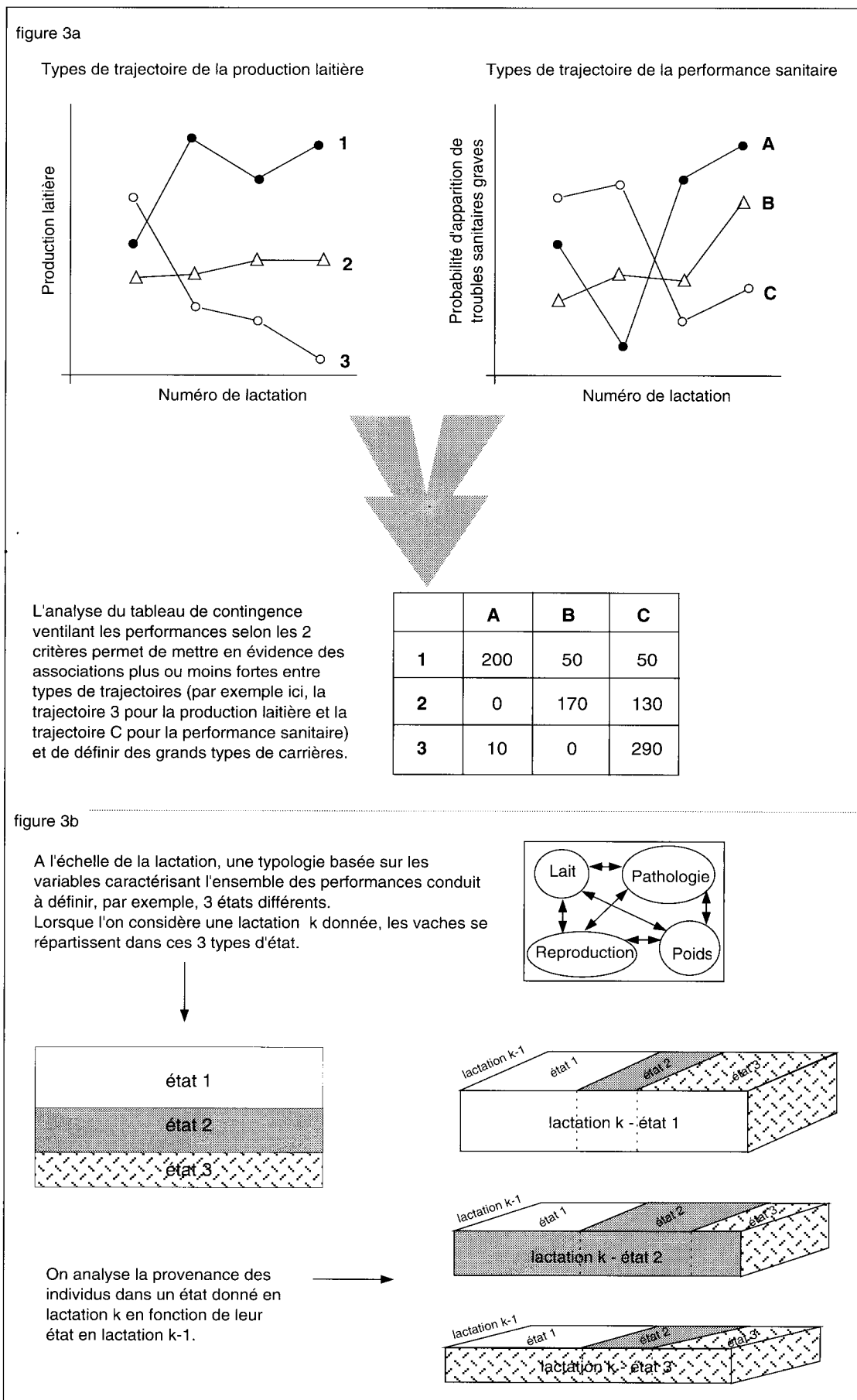




Figure 3. Deux conceptions d'étude de la carrière des vaches laitières : a) typologie de trajectoires par catégorie de performance et croisement des types, b) enchaînement d'états fondés sur l'ensemble des performances, avec accent sur les transitions.



appréhender correctement des performances aussi dissemblables qu'une production laitière, une variation de poids vif ou un intervalle entre 2 vêlages. L'échelle de la lactation semble être un bon compromis pour l'étude conjointe de ces différentes performances. A cette échelle, un premier travail consiste à analyser chacune de ces performances (élaboration de profils sanitaires, de production, de reproduction...) et leurs interrelations (association santé-reproduction, santé-production...). Ensuite, l'étude de la carrière peut se concevoir de différentes manières (figure 3) :

- La carrière peut être vue comme un ensemble de trajectoires correspondant chacune à une catégorie de performance (production laitière, reproduction, santé, variation de poids). L'objectif est de définir des types de trajectoires pour chaque catégorie de performance et de préciser l'existence éventuelle d'associations privilégiées entre les types de trajectoire de 2 catégories de performance. Un exemple théorique de cette démarche mettant en jeu 2 catégories de performance (production

laitière et troubles sanitaires) est présenté sur la figure 3a.

- La carrière peut être considérée comme un enchaînement d'états définis à l'échelle d'une lactation par l'ensemble des performances (Coulon *et al* 1989). L'objectif est d'analyser cet enchaînement (figure 3b).

- Enfin, on peut choisir a priori des descripteurs synthétiques de la carrière (poids au premier vêlage, production moyenne par jour de vie productive, nombre moyen d'inséminations par lactation, etc...), et définir des types de carrières à partir de ces descripteurs, comme cela a déjà été entrepris chez la brebis (Lasseur et Landais 1992).

Parallèlement, LASCAR peut permettre l'étude de problèmes importants se situant à une échelle de temps réduite (par exemple : effet à court terme des principaux troubles sanitaires sur la production laitière). Par ailleurs des questions plus marginales peuvent aussi être abordées (par exemple : analyse de la variation des heures de vêlage ou du poids des veaux).

## Références bibliographiques

- Barnouin J., Karaman Z., 1986. Enquête écopathologique continue : 9. Influence du niveau de production sur la pathologie de la vache laitière. *Ann. Rech. Vét.*, 17, 331-346.
- Broster W.H., Broster V.J., 1984. Reviews of the progress of dairy science : long-term effects of plane of nutrition on the performance of the dairy cow. *J. Dairy Res.*, 51, 149-196.
- Coulon J.B., Landais E., Garel J.P., 1989. Pathologie et productivité de la vache laitière : interrelations à l'échelle de la lactation. *Ann. Rech. Vét.*, 20, 443-459.
- Coulon J.B., Landais E., Garel J.P., 1990. Alimentation, pathologie et productivité de la vache laitière : interrelations à l'échelle de la carrière. *Ann. Rech. Vét.*, 21, 33-47.
- Dohoo I.R., Martin S.W., Meek A.M., Sandals W.C.D., 1982-1983. Disease, production and culling in holstein-friesian cows. I. The data. *Prev. Vet. Med.*, 1, 321-334.
- Enevoldsen C., Grohn Y.T., Thysen I., 1991. Sole ulcers in dairy cattle : associations with season, cow characteristics, disease, and production. *J. Dairy Sci.*, 74, 1284-1298.
- Ganière J.P., André-Fontaine G., Drouin P., Faye B., Madec F., Rosner G., Fourichon C., Wang B., Tillon J.P., 1991. L'écopathologie : une méthode d'approche de la santé en élevage. *INRA Prod. Anim.*, 4, 247-256.
- Hall S.A., 1978. Farm animal diseases data banks. *Adv. Vet. Sci. Comp. Med.*, 22, 265-286.
- Institut de l'Élevage-INRA, 1992. Résultats des contrôles de descendance des taureaux Prim Holstein. Production laitière, format et conformation. Edition 92/1. Ed. Institut de l'Élevage.
- Institut de l'Élevage-INRA, 1992. Résultats des contrôles de descendance des taureaux utilisés sur femelles montbéliardes. Production laitière, format et conformation. Edition 92/1. Ed. Institut de l'Élevage.
- Lasseur J., Landais E., 1992. Mieux valoriser l'information contenue dans les carnets d'agnelage pour évaluer des performances et des carrières de reproduction en élevage ovin-viande. *INRA Prod. Anim.*, 5, 43-58.
- Lescouret F., Pérochon L., Coulon J.B., Faye B., Landais E., 1992. Modelling an information system using the MERISE method for agricultural research : the example of a database for a study on performances in dairy cows. *Agric. Systems*, 38, 149-173.
- Lucey S., Rowlands G.J., Russel A.M., 1986. Short-term associations between disease and milk yield of dairy cows. *J. Dairy Res.*, 53, 7-15.
- Pérochon L., 1990. Modélisation par la méthode MERISE d'un système d'informations zootechniques et sanitaires à usage scientifique. Mémoire de fin d'études. INRA-CUST Clermont-Fd.
- Rolland C., 1986. Introduction à la conception de systèmes d'information et panorama des méthodes disponibles. *Génie Logiciel*, 4, 6-11.
- Tardieu H., Rochfeld A., Coletti R., 1983. La méthode MERISE. Tome 1. Principe et outils. Les Editions d'Organisation, Paris.

Udomprasert P., Williamson N.B., 1990. The dairychamp program : a computerised recording system for dairy herds. *Vet. Rec.*, 127, 256-262.

Wiktorsson H., 1979. General plane of nutrition for dairy cows. In : *Feeding strategy for the high yielding dairy cow*, Broster W.H et Swan H. Eds. Granada, St Albans, UK, 148-170.

Williams N.T., 1986. Using database management systems in farm management. *Farm Manage.*, 6, 183-191.

#### Remerciements

---

Nous tenons à remercier l'ensemble des personnels des domaines de Theix, Orcival et Marcenat qui ont réalisé la collecte des données, et en particulier, A. Ollier, J. Rouel et J.P. Garel, ainsi que les personnes des services communs du centre de Theix qui ont participé à l'informatisation des données (C. Barreyre, P. Champciaux, C. Espinasse, P. L'Hotelier).

---

### Summary

---

*Description of LASCAR, a database for studying the productive life of the dairy cow.*

A relational database called LASCAR has been designed to study the productive life of dairy cows. Data was recorded over the last 20 years in the 3 experimental dairy farms of INRA's Centre of Clermont-Fd-Theix. It is composed of 10 tables describing the genetic characteristics and measurements at the heifer period of the animals, individual performance on different time scales (milk production, liveweight, reproductive events,

diseases) and a few factors governing these performances (nutritional supplies, climate). It contains 3851 lactations from 1179 cows. Its data were exhaustively checked. Different ways of studying the productive life from this database are presented.

COULON J-B., LESCOURRET Françoise, FAYE B., LANDAIS E., TROCCON J-L., PÉROCHON L., 1993. Description de la base de données "LASCAR", un outil pour l'étude des carrières des vaches laitières. *INRA Prod. Anim.*, 6 (2), 151-160.