

PROBLÈMES STATISTIQUES POSÉS PAR L'HARMONISATION DES MÉTHODES DE CLASSEMENT DES CARCASSES DE PORC DANS LA C.E.E.

G. DAUMAS (1), T. DHORNE (2)

(1) I.T.P. - Pôle Qualité du Produit, BP 3, 35650 Le Rheu.

(2) I.N.R.A. - Laboratoire de Biométrie, 65 Rue de St-Brieuc, 35042 Rennes.

Cet article expose les principaux problèmes statistiques relatifs à l'élaboration des méthodes de classement des carcasses de porc et suggère certains éléments de réponse. Les problèmes sont sériés selon 4 axes : le taux de muscle, l'appareil de classement, le modèle de prédiction et le protocole expérimental.

Le taux de muscle, référence CEE, doit être défini très précisément et ne laisser place à aucune interprétation possible. De plus, le coût expérimental de son obtention doit être modéré. Ceci évite de recourir à la méthode du double échantillonnage qui pose des problèmes statistiques trop importants.

Un gain de précision substantiel dans la prédiction peut être obtenu en utilisant un modèle dépendant de certains facteurs, notamment le facteur sexe, et non linéaire dans les prédictrices. Les erreurs de mesure de ces prédictrices, et notamment celle liée à la détection difficile de la limite entre la noix de côtelette et les intercostaux, doivent être prises en compte.

L'hypothèse vraisemblable d'indépendance entre l'erreur de mesure et le taux de muscle permet de tester chaque appareil sur un échantillon indépendant de celui de la dissection. Le modèle de prédiction peut donc être mis au point avec un appareil de référence très précis et les appareils industriels peuvent alors être comparés à celui-ci par un modèle mixte à 1 facteur (appareil). Le protocole expérimental inclura un contrôle de l'effet opérateur et une stratification sur les prédictrices.

Statistical problems encountered in harmonization of methods for grading pig carcasses in the E.E.C.

This paper deals with the main problems encountered in the development of pig carcass grading methods and some solutions to these. Problems are referenced from 4 points of view : lean meat percentage, grading probes, prediction model, experimental design.

The EC-reference lean meat percentage must be defined precisely in order to avoid any misinterpretation. Furthermore, it has to be obtained at a moderate cost. This avoids the use of double sampling which generates too many statistical problems.

Prediction can be largely improved by using different models according to some factors, especially sex, and non linear with respect to the explanatory variables. The measurement error made on these explanatory variables, particularly this due to the difficult appreciation of the limit of the eye muscle, must be taken into account.

Under the assumption of independence between measure error and lean meat percentage, it is possible to test any grading probe on a sample independent from the dissected sample. The prediction model can be based on one precise grading used as a reference ; then probes are compared to this reference using a one-way (probe) mixed model. The experimental design should control operator effect and be stratified on the explanatory variables.

L'ancien «classement subjectif» des carcasses de porc, a été remplacé par un «classement objectif», basé sur une estimation statistique du taux de muscle des carcasses. Cette nouvelle grille a été définie en 1984 (règlement (CEE) n° 3220/84) et ses modalités d'application en 1985 (règlement (CEE) n° 2967/85). Ce «classement objectif» est obligatoire dans tous les Etats membres de la C.E.E. depuis le 1er janvier 1989. Mais force est de constater que 3 ans plus tard un certain nombre de pays ne respectent toujours pas la réglementation.

Par ailleurs, la pratique des pays utilisant la classification au taux de muscle a révélé des distorsions. Ces dernières sont à attribuer à l'insuffisance de la réglementation, d'une part, et à un manque de rigueur dans l'examen des dossiers de demande d'homologation des méthodes, d'autre part. Aussi, la France a demandé à la Commission des CE de réaliser une harmonisation des méthodes de classement au plan communautaire. Devant l'intérêt d'un tel objectif, la Commission, avec l'aide de quelques experts européens, a défini un programme de recherche sur ce sujet. Ce programme s'est pour l'instant limité à la réalisation d'un essai en 1990 et 1991 sur la base d'un protocole identique pour tous les Etats membres (Commission des C.E. 1990).

Certains Etats membres ont émis des réserves sur les objectifs et les moyens restreints de cet essai. Des contre-propositions écrites ont été formulées par la délégation française (DAUMAS et DHORNE, 1990). Ceci n'a pas abouti à une modification du protocole.

Aussi, il apparaît indispensable de cerner l'intégralité des problèmes techniques et statistiques posés par l'élaboration de méthodes de classement des carcasses de porc. Tel est l'objectif de cet article. Ces problèmes sont abordés selon 4 axes : le calcul du taux de muscle, l'appareil de classement, le modèle de prédiction et le protocole expérimental. Pour certains points, des éléments de réponse sont proposés, soit sur la base de travaux antérieurs de divers auteurs, soit à l'initiative des auteurs.

1. DÉFINITION DU TAUX DE MUSCLE

La réglementation actuelle définit le taux de muscle (ou teneur en viande maigre) d'une carcasse de porc comme le rapport entre un poids de muscles et un poids de carcasse, exprimé en pour cent.

1.1. Poids des muscles

Le règlement (CEE) n° 3220/84 stipule que le poids de muscles correspond au poids de l'ensemble des muscles rouges striés obtenus par dissection totale de la carcasse, pour autant qu'ils puissent être séparés à l'aide d'un couteau. En pratique, on se réfère souvent à la «dissection totale de Kulmbach», méthode pratiquée par l'Institut Allemand de Recherches sur la Viande (BAFF), situé en Bavière, à Kulmbach. Or, les Allemands incluent le poids du rognon dans le poids des muscles, ce qui surestime le taux de muscle d'environ 0,4 pour cent.

Par ailleurs, une affectation différente de mélanges muscle-graisse, notamment au niveau de la gorge, peut également engendrer des distorsions.

Mais le problème fondamental réside dans le règlement (CEE) n° 2967/85 qui autorise au lieu d'une dissection totale une méthode nationale de dissection d'effet équivalent. Sans contraintes statistiques précises, ceci permet toutes les déri-

ves imaginables dans le calcul du poids des muscles. En particulier, lorsqu'une partie du poids de muscles est estimé, le calcul des propriétés statistiques des prédicteurs s'en trouve grandement compliqué.

1.2. Poids de la carcasse

Le règlement (CEE) n° 3220/84 définit la «carcasse de porc» comme «le corps d'un porc abattu, saigné et éviscéré, entier ou divisé par le milieu, sans la langue, les soies, les onglons et les organes génitaux, mais avec la panne, les rognons et le diaphragme». Il s'agit d'un poids froid, obtenu en déduisant 2,0 % du poids constaté à chaud au plus tard 45 minutes après que le porc ait été égorgé.

Mais, des dérogations sont autorisées concernant une présentation différente de la carcasse à la pesée et un taux différent de ressuage. Une présentation sans certaines parties (exemple : sans panne ni rognons ni diaphragme) va conduire à une estimation du poids de la carcasse, ce qui entraîne des difficultés dans l'estimation des paramètres statistiques. Aussi, il conviendrait d'adopter une présentation minimale de la carcasse, c'est à dire sans les parties qui sont enlevées au moins dans un Etat membre.

Signalons par ailleurs le problème de la définition anatomique du diaphragme ; au niveau communautaire, on ne sait pas vraiment si ce terme recouvre l'onglet, la hampe ou les deux.

1.3. Taux de muscle

En pratique, le taux de muscle n'est pas calculé comme le simple rapport entre le poids des muscles, défini au § 1.1., et le poids de la carcasse, défini au § 1.2. En effet, on ne dissèque jamais une carcasse entière, ce qui serait beaucoup trop lourd, mais toujours une demi-carcasse. Il est surprenant que la réglementation ne le mentionne pas.

Le problème de la dissymétrie droite-gauche (POMMERET et NAVEAU, 1978) n'a pas été abordé.

Par ailleurs, les 2 termes du rapport du taux de muscle ne sont pas homogènes. La définition actuelle inclut les pertes de découpe et de dissection. On pourrait très bien envisager de calculer le poids de la (demi) carcasse disséquée par sommation des poids des tissus isolés après dissection.

2. APPAREILS

Par appareil, nous entendons ici l'ensemble du système qui comprend un instrument de mesure, un micro-processeur, un logiciel d'analyse du signal et de calcul du taux de muscle et une imprimante. Nous traiterons uniquement ici du cas des appareils à sonde, qui sont de loin les plus répandus (DAUMAS, 1991-a). Leur utilisation peut générer 2 types d'erreurs : celles liées à l'appareil, et celles liées à la prise de mesure.

2.1. Erreurs liées à l'appareil

Les erreurs liées à l'appareil peuvent survenir lors de l'acquisition des données ou lors de leur interprétation.

2.1.1. Erreurs d'acquisition

L'erreur d'acquisition dépend notamment de la précision du capteur d'enfoncement, 0.2 à 0.5 mm selon les sondes, et de

la résolution optique difficilement quantifiable. Les sondes qui ne gèrent pas le niveau de réflectance ne permettent pas d'interprétations des valeurs dans un milieu donné. De plus, si la capacité du couple émission - réception diminue (dynamique trop faible), alors l'acquisition devient moins fiable.

Signalons aussi que l'acquisition des données peut être faussée par un entretien inadapté ou encore par un mauvais étalonnage ; cette dernière source d'erreur est évitée dans certains systèmes qui ne peuvent fonctionner si l'étalonnage est mauvais.

2.1.2. Erreurs d'interprétation du signal

La résolution optique des sondes actuellement sur le marché, difficilement quantifiable, ne permet pas de situer avec exactitude la limite entre la noix de côtelette et les muscles intercostaux. Des tests peuvent éventuellement permettre de corriger des valeurs aberrantes.

Un certain nombre d'anomalies, telles que les «viandes pisseuses», les hématomes, le décollement des tissus, la présence de veines de gras intramusculaire, peuvent générer également des erreurs d'interprétation.

Enfin, les arrondis des mesures et du taux de muscle peuvent dégrader la précision finale.

Au total, les erreurs d'interprétation du signal reflètent la précision de l'appareil, qui est fonction de sa résolution, de la performance du logiciel d'analyse du signal et des modalités de correction des aberrations.

Il en découle une possibilité de chiffrer les variabilités associées aux différentes mesures et notamment les covariances entre mesures. La détection de frontières entre tissus induit en effet des effets de covariation : une surestimation de l'un entraîne une sous-estimation de l'autre et réciproquement.

2.2. Erreurs liées à la prise de mesure

L'effet opérateur est sans conteste la source la plus importante d'erreur liée à la prise de mesure.

2.2.1. Effet opérateur

L'action de l'opérateur se traduit par :

- * le respect ou non du site d'entrée de la sonde, défini par une position verticale (exemple : entre les 3 et 4^{èmes} dernières côtes) et une distance par rapport à la fente (exemple : 6 centimètres) ;
- * l'écart à l'horizontalité de la piqure ;
- * l'angle (théoriquement 90°) par rapport au plan tangent à la surface externe de la couenne au point d'entrée de la sonde ;
- * le trajet des diodes émettrice et réceptrice situées à l'extrémité de la sonde lors du retrait de cette dernière ;
- * la vitesse de retrait de la sonde.

Pour limiter l'effet opérateur, il faut contrôler la façon de faire des opérateurs et quantifier leur effet à l'aide de statistiques par opérateur, puis modifier leur comportement dans la mesure du possible.

Il est clair qu'un gaucher risque d'avoir plus de difficultés à bien piquer, tous les appareils étant prévus pour des droitiers. La hauteur du podium par rapport à la taille de l'individu et à la position des sites a une influence certaine sur la capacité à bien

piquer, notamment horizontalement. Un podium réglable et un site unique limite les risques. Enfin, la qualité de la fente de la carcasse est importante pour bien situer la position d'entrée de la sonde par rapport à la ligne médiane.

2.2.2. Effet porc

La présence, selon le site, de la panne (mal décollée), du filet, du diaphragme, de reste de poumon, dans le coffre peut provoquer une mesure erronée de l'épaisseur totale de la carcasse. Ceci peut éventuellement engendrer une erreur sur l'épaisseur de muscle, selon le mode d'analyse du signal.

2.2.3. Effet environnement

Un excès d'éclairage au niveau de la cavité abdominale peut masquer la sortie de sonde et donc fausser la mesure de l'épaisseur totale, dans le cas des sondes qui n'utilisent pas la modulation de lumière.

3. MODÈLE DE PRÉDICTION

Le modèle de prédiction utilisé peut être défini dans un cadre théorique. Dans la pratique un certain nombre de problèmes se posent qui conduisent à modifier ces références théoriques en fonction des contraintes subies ou des objectifs choisis.

3.1. Le modèle de référence

Nous recensons ici l'ensemble des facteurs contribuant à la mise en place d'un modèle de prédiction.

On regroupe sous le terme de prédictrices les variables continues mesurées par opposition aux facteurs qualitatifs connus, mal connus ou inconnus.

3.1.1. Prédictrices

La réglementation indique que le taux de muscle doit être estimé à partir de mesures physiques d'une ou de plusieurs parties anatomiques de la carcasse de porc. Parmi les prédictrices qui ont déjà été utilisées, on note :

- * des épaisseurs internes de tissus adipeux ou musculaires,
- * des épaisseurs à la fente de tissus adipeux ou musculaires,
- * des paramètres de conformation (largeur minimale de la demi-carcasse dans la région lombaire, largeur maximale du jambon, angle de rebondi du jambon),
- * le poids de la carcasse.

Un essai allemand, commenté par BRANSCHIED et al. (1987) et par DEMPFLÉ et al. (1990), a mis en évidence la supériorité des mesures internes sur les mesures à la fente et les paramètres de conformation. Il est cependant fort probable que ces paramètres de conformation apportent une information conjointe aux mesures d'épaisseurs importante. Leur capacité à approcher le type génétique des animaux culards conduit à penser qu'ils seraient très utiles dans l'affinement des modèles actuels (voir § 3.1.2.) par exemple en cas d'utilisation d'un robot.

Toutefois la combinaison d'une épaisseur latérale de gras et d'une de muscle (variables faiblement corrélées) permet d'atteindre une bonne précision. Si l'on utilise plusieurs épaisseurs de gras, BRANSCHIED et al. (1987) recommandent de choisir des emplacements aussi distants que possible.

Le choix des sites doit tenir compte de la capacité de prédiction des variables, de leur facilité de repérage et de l'absence d'inconvénients techniques (perforation du filet, gêne du diaphragme,...).

La multiplicité des sites de mesures a un double inconvénient :

- * surcharge de travail en conditions industrielles,
- * colinéarité des prédictrices.

En revanche dans le cas d'une forte variabilité de la mesure, la répétition de la mesure d'une même épaisseur à 2 sites différents peut fortement améliorer la prédiction car elle diminue le biais d'estimation dû à l'erreur de mesure sur les prédictrices.

Le poids est un élément dont l'importance relative est reconnue dans la prédiction du taux de muscle. L'évolution des différents tissus dépendant de l'âge, celle-ci est liée au poids. Il conduit à différentes stratégies de prédiction :

- * stratification de l'échantillon selon le poids :
 - prise en compte du poids dans le modèle
 - non prise en compte du poids dans le modèle (cette stratégie est statistiquement et intuitivement erronée)
- * prise en compte du poids comme variable de prédiction linéaire
- * prise en compte du poids comme facteur structurant (cas de l'Italie : une équation par gamme de poids)

Les 2 dernières stratégies bien qu'intéressantes peuvent être largement améliorées. Il est fort probable en effet que le poids agisse plutôt sur les paramètres de l'équation comme le suggère la dernière stratégie, et donc qu'un modèle multiplicatif soit plus adapté. En revanche la stratégie italienne est trop rigide car elle conduit à des problèmes de cohérence aux bornes des 2 intervalles de poids.

3.1.2. Facteurs structurant la variabilité

Il s'agit des variables qualitatives qui jouent un rôle important dans la variabilité : sexe, type génétique, mode d'alimentation,... Certains de ces facteurs sont connus (sexe), d'autres le sont moins précisément. Trois stratégies sont alors utilisables.

- * échantillonnage aléatoire simple (facteurs ignorés)
- * échantillonnage stratifié sur les facteurs
- * modèles de prédiction différents selon les modalités des facteurs.

Les 2 premières stratégies sont en théorie équivalentes en espérance, cependant certains facteurs étant mal connus, les choix de stratification conduisent à des différences théoriques d'espérance. Le protocole d'essai d'harmonisation (Commission des C.E. 1990) demande une stratification sur les types génétiques en proportions égales à celles de la population nationale. Cette contrainte étant difficile à vérifier, beaucoup d'imprécisions et de manipulations peuvent en découler. Par ailleurs comme l'indiquent de manière approximative KÖRMENDY et al. (1991), il n'y a pas convergence vers une valeur limite des paramètres pour la simple raison que cette valeur n'existe pas. Il s'ensuit que l'erreur de prédiction ne tend pas vers 0 lorsque la taille de l'échantillon tend vers l'infini.

L'existence de biais entre les sous-populations conduit à penser que la troisième stratégie est la meilleure, elle n'est cependant pas encore utilisée. Sous une forme théorique

optimale, il conviendrait de mettre au point une équation dépendant des modalités des facteurs c'est-à-dire utiliser la régression factorielle. Ceci est possible et pourrait être facilement mis en oeuvre pour le sexe. Il est important que les modèles tenant compte de l'effet sexe se généralisent car le sexe des animaux peut être facilement connu sur les chaînes d'abattage. La variabilité résiduelle est fortement réduite avec cette information. Le problème du type génétique est plus complexe car la classification en type génétique est plus délicate et la connaissance du type génétique plus difficile au moment du classement. Cependant, comme nous l'avons évoqué plus haut, il est possible de tenir compte de mesures annexes, par exemple de conformation, pour apporter des informations sur le type génétique probable et d'en tenir compte dans les modèles de prédictions.

3.1.3. Choix de la modélisation

Les modèles de prédiction utilisés sont de complexité très variable. Quelques pays utilisent des modèles non linéaires dans les prédictrices comme le suggèrent KÖRMENDY et al. (1991), mais la plupart des modèles sont de simples modèles linéaires.

Modèle linéaire ou non linéaire dans les prédictrices -

L'utilisation de modèles strictement linéaires dans les variables prédictrices pose le problème essentiel des biais globaux ou locaux d'estimation. Il est en effet peu probable que les espérances conditionnelles du taux de muscle par rapport à des prédictrices soient linéaires compte tenu du phénomène étudié. Il semble donc nécessaire d'utiliser des modèles plus complexes propres à mieux rendre compte de la réalité et spécialement aux bornes du domaine d'étude. Comme cela est suggéré par KÖRMENDY et al. (1991), l'emploi de modèles quadratiques peut être intéressant. Contrairement à ce que ces auteurs prétendent, les problèmes numériques engendrés par la colinéarité ne posent aucune difficulté, il suffit d'utiliser une arithmétique adaptée. En revanche les problèmes statistiques peuvent être gênants dans la sélection des meilleurs sous-modèles de prédiction. Il importe donc dans la pratique d'étudier les courbes de réponses pour préselectionner visuellement les modèles les plus adaptés.

Modèle linéaire ou non linéaire dans les paramètres - Il peut être judicieux dès lors, bien que ceci n'ait jamais été réalisé jusqu'ici, d'utiliser des modèles liés à la connaissance biologique de la différenciation tissulaire. Ces modèles déterministes et souvent différentiels sont non linéaires dans les paramètres à estimer. Les méthodes d'estimation sont plus complexes mais surtout les propriétés des estimateurs sont beaucoup plus difficiles voire impossibles à obtenir. Ces modèles pourraient cependant être utilisés comme référence théorique puis approchés par des modèles linéaires adaptés.

3.2. Critères statistiques

Le modèle (ou les sous-modèles) étant choisi(s), les critères d'ajustement doivent être explicités afin d'estimer les paramètres du modèle de prédiction en vue de son utilisation. Le choix des critères peut être guidé par des contraintes réglementaires (ce qui est partiellement le cas actuellement) ou par des contraintes d'efficacité statistique (ce qui est aussi partiellement le cas actuellement).

3.2.1. Critères généraux

Les critères généraux sont ceux qui sont utilisés couramment

dans le cadre du modèle linéaire, nous présentons leur principe, leurs avantages et inconvénients.

R^2 - Il correspond au rapport entre la somme de carrés expliquée par le modèle de régression et la somme de carrés totale de la variable à prédire, c'est aussi le carré du coefficient de corrélation entre la variable à prédire et le prédicteur obtenu. Ce critère a l'avantage de la simplicité, mais n'a par ailleurs que des inconvénients. Il est maximal (=1) quand le nombre de paramètres du modèle est égal au nombre d'observations, ce qui n'a aucun intérêt prédictif. Il n'a par ailleurs pas grand sens en cas de planification (KÖRMENDY et al., 1991). Ce critère est actuellement retenu par le règlement CEE 2967/85, mais la commission des experts statisticiens s'est prononcé pour sa suppression (COOK et al., 1990).

RSD - Il s'agit de l'écart-type résiduel, racine carrée de la variance résiduelle. Si le modèle est exact, la variance résiduelle empirique est le meilleur estimateur non biaisé de la variance résiduelle théorique. Cependant, le modèle exact n'étant pas connu, cette propriété n'est pas toujours intéressante. En revanche, l'écart-type résiduel peut être utilisé comme critère d'ajustement. Il est en effet souhaitable de minimiser cet écart-type, or celui-ci ne diminue pas indéfiniment lorsque l'on augmente le nombre de prédicteurs. Il est d'ailleurs indéterminé lorsque le nombre de paramètres du modèle est égal au nombre d'observations car dans ce cas, son numérateur et son dénominateur sont tous deux nuls.

R^2 ajusté - Ce critère a été défini pour pallier l'inconvénient majeur du R^2 à savoir sa croissance monotone en fonction du nombre de prédictrices. Le R^2 ajusté est le rapport de la variance expliquée par le modèle à la variance totale. Les sommes de carrés sont donc divisées par les degrés de liberté associés ce qui explique le fait que ce critère passe par un maximum.

3.2.2. Critères liés à la prédiction

Cp de MALLOWS - Le Cp de MALLOWS est un critère utilisé pour choisir le meilleur modèle prédictif dans une famille de modèles dominés c'est-à-dire tous sous-modèles d'un modèle donné. L'existence d'un modèle dominant étant problématique on pourrait utiliser comme référence un modèle de lissage.

RSD - Lié au Cp de MALLOWS, le rsd est intéressant pour la prédiction. L'erreur de prédiction estimée est proportionnelle au rsd mais est aussi fonction des coordonnées des points de prédiction. La précision est d'autant moins bonne que l'on se rapproche des bornes du domaine. Il conviendrait donc de tenir compte de l'erreur de prédiction aux bornes du domaine d'utilisation des équations.

MSEP - Un certain nombre de travaux se sont posé la question de la minimisation de l'erreur de prédiction (Mean Square Error of Prediction), on peut citer en particulier HEBEL, 1989 qui l'a appliqué au cas de la prédiction des rendements de blé.

3.3. Estimation des paramètres et critères associés

Le type d'ajustement utilisé dépend fortement :

- * du type de modèle
- * du ou des critères d'ajustement retenu(s)
- * des hypothèses statistiques posées

Nous étudierons d'abord les problèmes généraux posés par

l'estimation puis nous verrons comment certaines conditions expérimentales conduisent à des hypothèses statistiques particulières posant des problèmes d'estimation.

3.3.1. Problèmes généraux

Le problème le plus important dans le choix de la stratégie d'ajustement est celui de la prise en compte des observations. En effet la réalisation de l'expérience peut conduire à des résultats sur lesquels il faut être très critique. L'étude détaillée des données obtenues dans plusieurs expériences a montré que la part de données erronées est assez importante (entre 5 et 15 %). Il est donc souhaitable de réaliser soit une estimation robuste des paramètres soit une analyse interactive des données. L'objectif est de ne pas donner une contribution trop forte à certaines observations dans le modèle de prédiction. Il nous paraît en effet préférable de mettre en place un modèle qui est optimal pour la majorité des observations et inadapté pour une petite proportion plutôt qu'un modèle intermédiaire mal adapté pour l'ensemble des observations. Il est alors possible de produire d'autres stratégies de prédiction pour certains cas extraordinaires.

Lorsque ces problèmes ont été réglés, il paraît naturel d'estimer les paramètres par la méthode statistique la plus adaptée au modèle utilisé c'est-à-dire le maximum de vraisemblance ou l'estimateur non biaisé de variance minimale qui sont confondus pour la plupart des modèles utilisés.

Cependant d'autres techniques sont utilisables et même peut être plus adaptées comme l'estimation bayésienne lorsque l'on connaît les caractéristiques des populations porcines.

Par ailleurs certaines conditions expérimentales particulières imposent des techniques d'estimation plus complexes.

3.3.2. Le modèle du double échantillonnage

Ce modèle est utilisé lorsque l'on cherche à accroître la précision ou à diminuer les coûts dans la mise au point de modèles de régression grâce au remplacement partiel de dissections totales coûteuses par des dissections partielles. Il a été largement étudié par CONNIFFE et MORAN (1972), COOK et al. (1983), ENGEL et WALSTRA (1991). L'intérêt économique manifeste de ce modèle s'accompagne cependant de beaucoup de restrictions éventuelles d'utilisation. Les problèmes statistiques posés sont en effet fort nombreux (ENGEL et WALSTRA, 1991), en particulier les propriétés connues sont asymptotiques et il est impossible de calculer les critères réglementaires à distance finie autrement que par simulation. Ce modèle semble donc difficile à mettre en pratique, une modification de la référence nous paraît plus judicieuse.

3.3.3. Le modèle avec erreur sur les variables explicatives

Le modèle de régression déterministe classique suppose que les variables prédictives sont connues sans erreur. Il est manifeste comme nous l'avons évoqué plus haut (cf. § 2.) que la mesure des prédictives est entachée d'une erreur qui peut parfois être importante. Contrairement à ce que prétendent certains auteurs, ce problème ne disparaît pas dans le cas de la régression probabiliste et les erreurs de mesure engendrent des biais sur les estimations des paramètres et ceci d'autant plus forts que les variances des erreurs sont élevées. Il semble donc important, dès lors qu'une estimation de ces variances est rendue possible, que les modèles de prédiction en tiennent compte.

3.3.4. Le modèle avec test séparé des appareils

Les remarques précédentes conduisent, contrairement à ce que supposent les règlements communautaires, à penser qu'il existe une dissociation entre la mesure des prédictrices et la mise au point d'une équation de prédiction du taux de muscle. Cette idée a été mise en pratique en France en 1987 puisqu'à la différence de ce qui se faisait dans les autres pays de la Communauté il n'existe dans notre pays qu'une seule équation de prédiction quelque soit l'appareil utilisé. L'Allemagne a depuis, suivi et prolongé l'initiative française en proposant un cadre formel à cette pratique. Les aspects opérationnels ont été étudiés en Allemagne (ANONYME, 1989) et les aspects statistiques par DEMPFLÉ (1989). Sous des hypothèses acceptables dans la pratique, DEMPFLÉ (1989) détaille les procédures de calcul des critères statistiques retenus pour valider un modèle de prédiction. Bien que cette pratique conforte le problème évoqué précédemment (§ 3.3.3.) des erreurs de mesure, celui-ci n'est cependant pas abordé conjointement.

3.3.5. Le modèle de prédiction à risque quadratique minimal

Le choix d'un modèle de prédiction à risque quadratique minimal suggéré plus haut (cf. § 3.2.2.) pose le problème de la procédure d'estimation associée. Pour plus de détails sur cette procédure, nous renvoyons le lecteur à HEBEL (1989).

4. PROTOCOLE EXPÉRIMENTAL

Le choix d'un modèle statistique particulier constitue l'objectif d'une expérience particulière. Le protocole expérimental correspond alors aux moyens pour atteindre au mieux cet objectif. Nous détaillons dans la suite les principes généraux et les protocoles expérimentaux possibles en fonction des choix opératoires réalisés.

4.1. Principes généraux

Il existe schématiquement 2 types de protocoles expérimentaux :

- * les protocoles aléatoires (sur les prédictrices)
- * les protocoles planifiés (stratifiés sur les prédictrices)

Nous ne parlons pas ici des facteurs pris en compte dans le modèle qui ont été évoqués plus haut, mais uniquement des prédictrices.

Les protocoles planifiés nécessitent la mesure de variables avant l'échantillonnage, ce qui peut poser des problèmes sur la chaîne d'abattage, mais ont l'avantage, à taux d'échantillonnage comparable, de fournir de meilleures estimations des paramètres. C'est la technique qui a été utilisée pour l'expérimentation française de 1990 (DAUMAS, 1991-b). Il importe en général pour les modèles linéaires de n'utiliser que des observations ayant des valeurs extrêmes pour les prédictrices. Cependant dans le cas de modèles non linéaires dans les variables ou les paramètres un échantillonnage régulier peut être préconisé. Dans les 2 cas, il convient de remarquer que les proportions dans l'échantillon ne sont pas les proportions dans la population ce qui montre que les protocoles aléatoires sont moins adaptés. Une solution intermédiaire entre protocole aléatoire et protocole planifié consiste à structurer sur une combinaison linéaire des prédictrices, par exemple le taux de muscle estimé.

4.2. Cas d'un seul appareil

Lorsque l'on n'utilise qu'un seul appareil de mesure des prédictrices pour un site donné, il convient de définir de manière optimale le site en question. Notre point de vue est que la variabilité locale de la valeur théorique des mesures ne peut avoir d'incidence sur la prédiction. Aucun bon argument statistique ne peut conduire à penser qu'une épaisseur mesurée 0.5 cm plus haut ou plus bas qu'une autre va donner un prédicteur différent (statistiquement). En effet, la continuité du phénomène étudié plaide en faveur d'une covariance spatiale entre toutes les mesures. La variabilité locale (veines de graisse, intercostal flou,...) n'ont pas de raison particulière d'apparaître à un site plutôt que 0.5 cm plus haut ou plus bas. En revanche, il est évident que certains sites sont plus propres à la prise de mesures que d'autres.

Il importe donc de spécifier clairement les sites de mesure ainsi que l'erreur tolérée (cf. § 2.2.1.) autour de celui-ci, ce qui n'est jamais fait aujourd'hui. Dans ce cas, l'estimation de la variabilité locale par plusieurs prises de mesures y compris au même site, nous paraît être une garantie de sérieux de l'expérience et contribue à une meilleure prédiction (cf. § 3.3.3.).

4.3. Cas de plusieurs appareils

4.3.1. Plusieurs équations de prédiction

Lorsque plusieurs appareils sont utilisés et qu'à chacun d'entre eux correspond une équation de prédiction, nous nous trouvons dans le même cas que précédemment pour chacun des appareils lorsqu'une expérience est mise en place pour chaque cas.

Cependant, très souvent, pour des raisons d'économie, une seule expérience est mise en place pour tous les appareils. Il s'ensuit que les conditions de mesure définies plus haut ne sont pas les mêmes pour chaque appareil ou alors que dans la pratique on s'autorise un certain nombre d'écarts (sites flous, possibilité de repiquer plusieurs fois au même endroit,...). Quelque soit la solution choisie, il est évident qu'il existe une erreur supplémentaire à celle que l'on observe dans le cas d'un seul appareil (erreur de décalage, erreur de repiquage,...). Il importe donc qu'il en soit tenu compte aussi bien dans le modèle (cf. § 3.3.3.) que dans le protocole (randomisation).

Cette stratégie nous paraît être la plus discutable du point de vue de la cohérence. Il paraît en effet difficile que l'équation de prédiction théorique (c'est-à-dire en l'absence d'erreur de mesure) ne soit pas la même pour tous les appareils. Dans ce cas les équations pour les différents appareils ne doivent différer que statistiquement si elles sont construites à partir d'échantillons différents. En revanche si elles sont élaborées à partir du même échantillon, il paraît plus judicieux de considérer la mise au point de l'équation et le test des appareils comme deux étapes indépendantes, ce qui est exposé dans le paragraphe suivant.

4.3.2. Une équation de prédiction

On suppose dans ce cas, ce qui paraît assez intuitif, que l'appareil n'est là que pour mesurer une prédictrice. La connaissance de cette prédictrice permet alors d'élaborer un modèle statistique de prédiction. Ces deux étapes sont donc séparées, ce qui se traduit statistiquement par le fait que ou bien il n'y a pas d'erreur de mesure, ou bien l'erreur de mesure est indépendante de la variable à prédire. C'est l'hypothèse qui

est faite dans la suite et qui correspond à celle de DEMPFLÉ et al. (1990).

Il convient alors de choisir un appareil de référence, utilisé dans la mesure des prédictrices lors de la mise au point de l'équation. Cet appareil doit être si possible le plus précis des appareils existants, car il permet alors la meilleure fiabilité dans la mise au point de l'équation. L'estimation des paramètres de l'équation doit cependant tenir compte de l'erreur de mesure.

Par la suite, si pour des raisons particulières (coûts, cadences,...) on souhaite utiliser un autre appareil, il convient de le comparer à l'appareil de référence. On impose alors qu'il soit non biaisé (au besoin par une correction), puis on étudie sa variance d'erreur dont il est tenu compte lors du calcul des critères statistiques lié à l'équation de prédiction. Le non biais est fondamental pour que toute carcasse soit jugée en espérance de la même façon avec différents appareils et donc qu'il n'y ait aucun avantage systématique à être classé par un appareil plutôt que par un autre.

Lorsque ces conditions sont vérifiées la mise au point de l'équation se fait alors dans les mêmes conditions que précédemment. En revanche le test des appareils doit se faire selon un protocole particulier mentionné dans la suite.

4.4. Cas du test séparé des appareils

Le test séparé des appareils nécessite la vérification du non

biais par rapport à un appareil de référence supposé non biaisé de la vraie valeur ainsi qu'une estimation de l'erreur de mesure de l'appareil à tester. Il convient de noter que plus l'appareil de référence est précis, moins les appareils à tester n'auront à être précis pour que les critères globaux d'homologation soient vérifiés. En effet la contrainte liée au RSD correspond à une somme entre la variance de l'appareil à tester et de la variance de l'appareil de référence. Plus l'une est faible plus l'autre peut être élevée.

Cependant pour éviter l'homologation d'appareils peu fiables, certains pays comme la France ou l'Allemagne ont imposé des critères plus sévères pour le test des appareils indépendamment de l'équation de prédiction.

Le protocole proposé pour tester un appareil peut donc être conçu comme un modèle mixte à 1 facteur appareil à 2 modalités (référence et à tester). D'autres facteurs peuvent être ajoutés au modèle s'ils ne peuvent être randomisés (opérateur, piquage-repiquage, site,...). Notre expérience nous fait penser que ces facteurs sont assez peu influents en conditions expérimentales optimales. En particulier, nous avons montré que le fait de repiquer au même site n'entraîne aucun biais au niveau du signal (DAUMAS et al. 1991).

REMERCIEMENTS

à l'OFIVAL pour sa contribution.

RÉFÉRENCES BIBLIOGRAPHIQUES

- ANONYME, 1989. Communication des méthodes de classement devant être utilisées en Allemagne. Document de travail CEE VI/2141/89.
- BRANSCHIED W., KOMENDER P., OSTER A., SACK E., FEWSON D., 1987. Züchtungskunde 59, (4), 258-280.
- Commission des C.E., 1985 Règlement CEE n° 2967/85 du 24 Octobre 1985 établissant les modalités d'application de la grille communautaire de classement des carcasses de porcs - JOCE n° L285 du 25.10.85.
- Commission des C.E., 1990. Recherche concernant l'harmonisation des méthodes de classement des carcasses de porc dans la Communauté. Document (CEE) VI/3860/89 Rév. 6.
- CONNIFFE D., MORAN M.A., 1972. Biometrics 28, 1011-1023.
- Conseil des C.E., 1984. Règlement CEE n° 3220/84 du 13 Novembre 1984 déterminant la grille communautaire de classement des carcasses de porcs - JOCE n° L301 du 20.11.84.
- COOK G.L., JONES D.W., KEMPSTER A.J., 1983. Anim. Prod. 36, 493-495.
- COOK G.L., DHORNE T., ENGEL B., 1990. Propositions de modification des critères statistiques utilisés pour l'homologation des modèles de prédiction du taux de muscle. Note interne CEE.
- DAUMAS G., 1991-a. Clasificación de las canales porcinas en la CEE. Jornadas Científicas SEPOR, 7-12.
- DAUMAS G., 1991-b. Recherche concernant l'harmonisation des méthodes de classement des carcasses de porc dans la Communauté : compte-rendu national de l'essai réalisé en France.
- DAUMAS G., DHORNE T., 1990. Remarques sur le contenu du projet de recherche concernant l'harmonisation des méthodes de classement des carcasses de porc dans la Communauté. Document de travail (CEE) VI/2241/90.
- DAUMAS G., DHORNE T., JOLY B., 1991. Effet induit par un repiquage au même site en classement de carcasse de porc. Compte rendu d'expérience.
- DEMPFLÉ L., 1989. Comments on the calibration procedure. Communication personnelle (non publié).
- DEMPFLÉ L., BRANSCHIED W., GRÜNDL E., SACK E., 1990. Analyse de l'essai de dissection et de classement des carcasses de porcs dans le cadre de l'harmonisation des sites de mesure. Document de travail (CEE) VI/524/90.
- DHORNE T., 1990. Protocole de comparaison d'appareils de classement. Document interne du Laboratoire de Biométrie INRA de Rennes.
- ENGEL B., WALSTRA P., 1991. Biometrics 47, 13-20.
- HALLÉPÉE F., 1990. Elaboration d'une méthode de suivi des classements de porcs charcutiers. Mémoire d'Ingénieur de l'ENITA de Dijon.
- POMMERET P., NAVEAU J., 1978. Composition anatomique de la carcasse : la Découpe Parisienne Normalisée. Rapport ITP 79-06.
- HEBEL P., 1989. Essai de Modélisation de la prévision des rendements nationaux de blé tendre d'hiver. Rapport technique du Département de Biométrie de l'INRA.
- KÖRMENDY L., CSIBA A., KLOSZ T., 1991. Viandes Prod. Carnés Vol. 12 (1) 3-6.