

Détermination de la composition de la carcasse de porc et la composition des coupes principales selon différentes méthodes d'analyses statistiques

Maria FONT-I-FURNOLS (1), Félix-Antoine OUELLET (2), Hugo LAROCHELLE (2), Marcel MARCOUX (3), Candido POMAR (3)

(1) IRTA-Qualité du Produit, Finca Camps i Armet, 17121 Monells, Catalogne, Espagne

(2) Université de Sherbrooke-Département d'informatique, 2500 Boulevard de l'Université, Sherbrooke, Québec, Canada, J1K 2R1

(3) AAC-Centre de Recherche et de développement sur le bovin laitier et le porc, 2000 rue Collège, Sherbrooke, Québec, Canada, J1M 0C8

Candido.Pomar@agr.gc.ca

Ce travail a été réalisé grâce au financement accordé à Mme Maria FONT-I-FURNOLS de OECD Programme Coopératif de Recherche : Biological Resource Management for Sustainable Agricultural Systems pour 2012-13.

Comparing different statistical approaches to determine carcass and cut composition

Carcass lean yield (RM) is the most important carcass quality parameter used in carcass evaluation. However it is also of interest to determine weight and composition of the main cuts to estimate carcass value. Subcutaneous fat and muscle depths are the most common variables used to predict carcass RM, while multiple linear regression is used to develop prediction equations. The aim of this study was to develop and compare multiple linear regression (LR), k-nearest neighbour (kNN) and Gaussian process (GP) statistical methods, when used to predict carcass RM, and weight and lean, fat and bone composition of the 4 main cuts from backfat and muscle depths. All the statistical methods used in this study make it possible to obtain an acceptable precision in the prediction of the variables of interest. Nevertheless, those obtained with LR and GP produced the lowest errors of prediction in comparison with kNN. Regarding the variables of interest, the best results were those obtained when estimating carcass RM as well as the proportion of lean in the shoulder, ham and loin. Thus, using only one fat thickness and one muscle depth as predictors, LR is the most simple and precise method to predict the different composition characteristics of carcasses and main cuts.

INTRODUCTION

Le rendement en maigre (RM) est le critère de qualité le plus utilisé dans les systèmes de classification de carcasses de porc (Pomar *et al.*, 2008). Toutefois, le poids et la composition des coupes principales sont essentiels pour estimer la valeur commerciale de la carcasse (Marcoux *et al.*, 2007). L'épaisseur du gras et du muscle dorsal sont les variables utilisées généralement pour prédire le RM à l'aide d'équations de régression (LR). Toutefois, la méthode des k plus proches voisins (kNN) et le processus gaussien (GP) peuvent aussi être exploités pour prédire le RM, car ils permettent d'établir les relations entre plusieurs variables d'intérêt simultanément. L'objectif de cette étude était d'évaluer et de comparer toutes ces méthodes d'estimation du RM de la carcasse ainsi que le poids et la teneur en maigre, gras et os des 4 coupes principales à partir de l'épaisseur du gras et du muscle dorsal.

1. MATERIEL ET METHODES

1.1. Carcasses et dissection

Soixante carcasses (29 femelles et 31 castrats) ont été utilisées. Les carcasses retenues étaient représentatives des

différents types de carcasses de porc observés dans les abattoirs selon le poids, l'épaisseur de gras et la conformation. Les épaisseurs du gras et du muscle ont été mesurées à 7 cm de la ligne médiane de la carcasse entre la 3^{ème} et la 4^{ème} dernières côtes avec la sonde Destron PG-100. Le jambon, la longe, le flanc et l'épaule ont été coupés selon la présentation standard canadienne avec un ajustement proportionnel à la longueur de la carcasse. Les quantités de muscle, de gras et d'os des coupes ont été obtenues par dissection. En plus du RM de la carcasse, les proportions de chaque tissu ont été calculées par rapport au poids de la coupe correspondante.

1.2. Analyses des données

Les approches statistiques suivantes ont été étudiées en utilisant les épaisseurs de gras et du muscle obtenues avec la sonde Destron PG-100 comme variables indépendantes. L'erreur de prédiction quadratique moyenne (RMSEP) et le coefficient de variation ($CV=100 \times RMSEP / \text{moyenne}$) ont été calculés à l'aide d'une validation croisée de type « leave-one-out ».

1.2.1. LR

La procédure REG du SAS (SAS Institute Inc., Cary, NC, USA) a été utilisée pour obtenir les équations de régression prédisant séparément chacune des variables d'intérêt.

1.2.2. *kNN*

La sélection de la valeur de k a été réalisée à l'aide d'une validation croisée de type « holdout », en isolant 15% de l'ensemble d'entraînement pour validation.

La valeur de k sélectionnée est alors celle minimisant l'erreur de prédiction sur ce sous-ensemble de validation.

Toutes les variables d'intérêt sont prédites avec cette méthode simultanément (multitâche).

1.2.3. *GP*

Un noyau linéaire a été choisi dans l'application d'un processus gaussien en mode multitâche, puisqu'une validation croisée a identifié que d'autres noyaux non-linéaires causaient des problèmes de surapprentissage.

Les valeurs des hyper-paramètres du noyau linéaire ont été obtenues par l'optimisation de la vraisemblance marginale des données d'apprentissage.

2. RESULTATS ET DISCUSSION

Lors de la prédiction du RM, le CV était respectivement de 3.48%, 5.25% et 3.60% pour les méthodes LR, *kNN* et GP. Toutes les méthodes utilisées ont permis d'obtenir une précision de prédiction acceptable des variables d'intérêt, quoique les méthodes LR et GP sont celles ayant une RMSEP relative (CV) entre 1.5% et 7.8% inférieure à celui obtenu avec la méthode *kNN* en dépendant de la variable prédite (Figure 1). Cependant, les meilleurs résultats ont été obtenus lors de la prédiction du RM de la carcasse et la proportion de muscle de l'épaule du jambon et de la longe.

CONCLUSION

Dans le contexte d'utilisation de seulement deux variables prédictives pour estimer le RM de la carcasse ou la composition de ses coupes principales, la méthode LR demeure la méthode la plus simple et la plus précise.

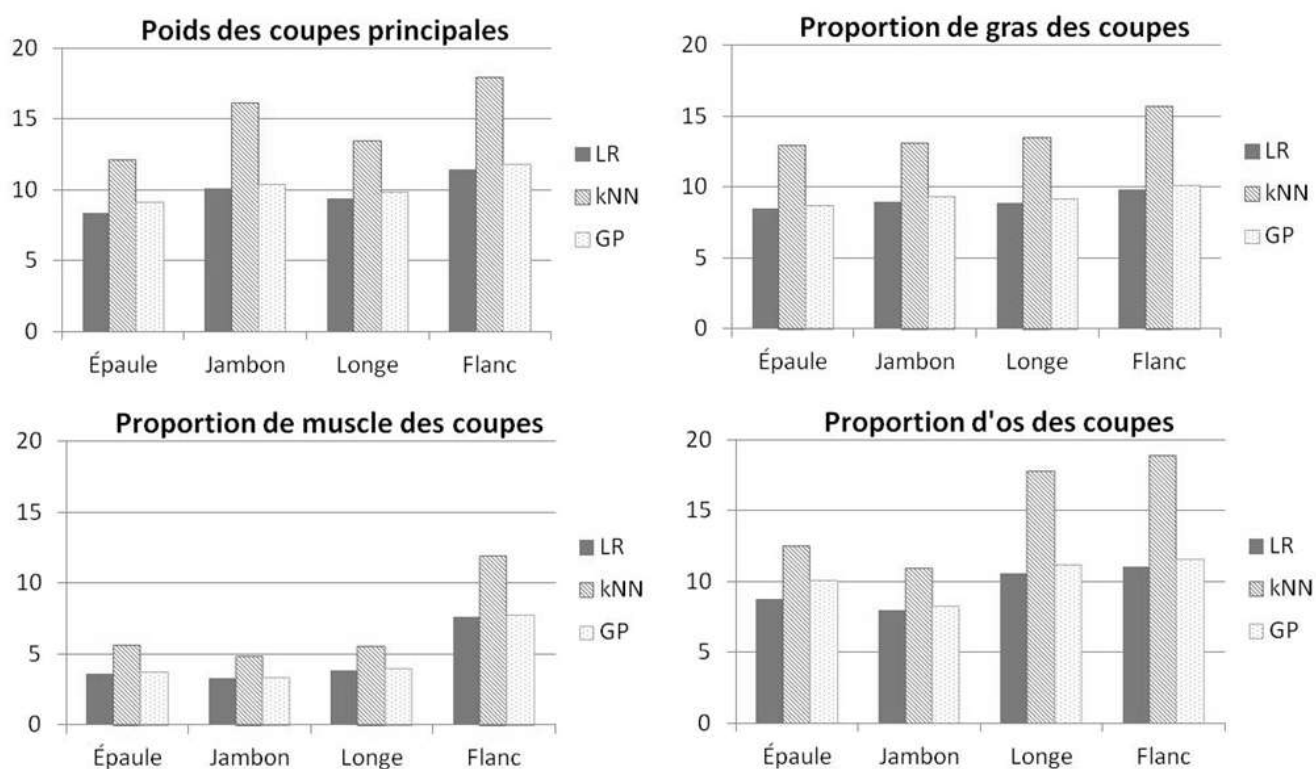


Figure 1 – Erreur de prédiction relative (CV(%)=100*RMSEP/moyenne) obtenu par régression linéaire (LR), k plus proches voisins (kNN) et le processus gaussiens (GP).

REFERENCES BIBLIOGRAPHIQUES

- Marcoux, M., Pomar, C., Fautitano, L., Brodeur, C. (2007). The relationship between different pork carcass lean yield definition and the market carcass value. *Meat Science*, 75, 94-102.
- Pomar, C., Marcoux, M., Gispert, M., Font i Furnols, M., Daumas, G. (2008). Determining the lean meat content of pork carcasses. In: *Improving the sensory and nutritional quality of fresh meat (Chapter 21)*. Ed. J. Kerry and D. Ledward. Woodhead Publishing Limited, Cambridge, UK (pp. 493-518).