

PATAPI : un outil de planification des pesées de bovins développé avec R

M. Legris * A. Delpeuch †

Résumé

Le projet PATAPI s'inscrit dans le cadre de l'amélioration de la gestion des pesées des bovins dans les élevages français, un domaine clé pour la performance génétique et la gestion optimale des troupeaux. Ce projet vise à fournir un outil performant de planification des pesées, en utilisant des algorithmes d'optimisation développés principalement en langage R. Le package Rcpp a été utilisé pour intégrer du code C++, permettant d'accélérer les calculs, et le package GenSA a été utilisé pour trouver des solutions optimales grâce à la méthode du recuit simulé. Une application Shiny a été créée pour aider au développement de l'algorithme en facilitant la visualisation des résultats et en permettant des ajustements dynamiques des paramètres. Le moteur de calcul a été encapsulé dans un package R et il est en cours d'APIfication. Il permettra ainsi de générer des prévisions sur les dates de pesée idéales, offrant un outil efficace et réactif aux techniciens et éleveurs, dans le but de simplifier la planification des pesées et d'améliorer la gestion des élevages.

Mots-clefs : Préviation - Optimisation - Package - Shiny

Développement

Contexte et objectif

En France, dans le cadre du Contrôle des Performances Viande (CPV), environ 400 000 poids à âge type (i.e. poids à 120 jours ou à 210 jours de vie de l'animal) sont calculés chaque année pour un peu plus d'un million de pesées, dont 80 % sont réalisées par des techniciens. Ces pesées sont utilisées pour conseiller l'éleveur dans sa conduite et alimentent l'évaluation génétique pour le calcul des index de croissance et de valeur laitière. Dans ce contexte, l'objectif du projet PATAPI est de développer un moteur de calcul permettant de planifier efficacement les visites des techniciens en élevage pour effectuer les pesées.

Les fonctionnalités de l'outil

Un premier travail préalable au projet PATAPI (Delpeuch (2021)) avait permis de définir des niveaux de fiabilité des estimations des poids à 120 jours et 210 jours en fonction de la position des dates de pesées dans la carrière de l'animal (interpolation/extrapolation, proximité des 120 jours ou 210 jours de l'animal...). L'outil développé dans le cadre de PATAPI devait proposer automatiquement des dates de pesées en élevage qui permettent d'estimer les poids des animaux à 120 et 210 jours avec un niveau de fiabilité suffisamment élevé. L'enjeu est de trouver un équilibre entre nombre de passage des techniciens en élevage et fiabilité de l'estimation des poids. Le moteur de calcul se base sur les dates de naissance des animaux (estimées ou réelles) et sur un ensemble de paramètres soit de performances souhaitées par l'éleveur/le technicien (ex : maximisation du poids à 120 jours, 210 jours ou les deux ; classes minimales de fiabilité...), soit liés aux contraintes de l'éleveur (ex : ne pas tenir compte d'animaux atypiques, définir des plages de dates interdites durant laquelle l'éleveur n'est pas disposé à peser, ne pas peser les animaux au-delà de l'âge moyen de sortie des animaux, fixer une date particulière de pesée...). L'outil final devait être à la fois un outil prévisionnel, i.e. qui permet d'avoir une simulation prévisionnelle des dates de passages en début de campagne, mais également un outil actualisé au fil de l'eau, i.e. qui permet d'ajuster les planifications en fonction des naissances réelles de l'élevage et en tenant compte des dates de pesées déjà réalisées.

*Institut de l'élevage, maxime.legris@idele.fr

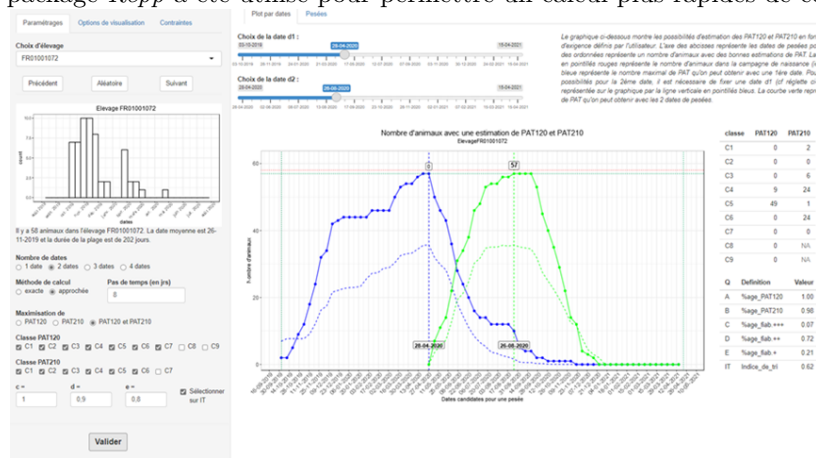
†Institut de l'élevage, arnaud.delpeuch@idele.fr

La fonction de coût et son optimisation via le package GenSA

Le travail dans le projet PATAPI consistait donc en la définition d'une fonction de coût, en son optimisation et en son intégration dans un dans un moteur de calcul robuste et unifié qui prend la forme d'un package R. Une premier travail a onstisté à réunir des experts du domaine pour définir une fonction de coût permettant de départager les solutions. Cette fonction de coût combine la proportion de poids bien estimés, le niveau de fiabilité de ces PAT et le nombre de visites nécessaires. Cette fonction de coût fournit un indice exprimé entre 0 (faible intérêt) et 1 (fort intérêt). La fonction de coût ainsi définie est une fonction avec de nombreux paramètres et contraintes qui définissent un espace de solution à plus de 3 dimensions, difficile à représenter graphiquement, et avec potentiellement de nombreux maxima locaux. Du fait de la complexité de cette fonction de coût, les techniques d'optimisation classique (descente de gradient, méthode de Newton...) ont été écartées au profit d'une méthode stochastique : le recuit simulé généralisé. Le package *GenSA* a été utilisé pour sa performance (Kirkpatrick, Gelatt, and Vecchi (1983), Xiang et al. (2013)).

L'application Shiny comme un démonstrateur

Tout au long du développement du moteur de calcul, l'utilisation d'une application *Shiny* a été une aide au développement de l'algorithme en facilitant la visualisation des résultats et en permettant des ajustements dynamiques des paramètres. Elle a permis de faire préciser le besoin des différents utilisateurs finaux de l'outil et a orienté le développement du moteur de calcul final vers deux modes de calcul : un mode *simplifié* avec uniquement un ensemble de dates de pesées optimales et un mode *complet* qui intègre en plus des courbes 2D. Le package *Rcpp* a été utilisé pour permettre un calcul plus rapides de ces courbes.



Un package R qui optimise les pesées

Le package R contenant le moteur de calcul développé dans PATAPI permet au final de proposer un ensemble de dates optimales pour les pesées et des résultats plus précis pour un pilotage des visites en élevage. Sur un échantillon d'élevages aux caractéristiques diverses (nombre d'animaux, étendue de la période de naissance), le moteur de calcul permet de réduire le nombre de visites et/ou d'améliorer le positionnement des dates de pesées. Les temps de calcul en mode simplifié sont raisonnables (inférieurs à 5 secondes pour la majorité des élevages) mais largement perfectibles et plusieurs pistes sont envisagées pour une deuxième version.

Références

- Delpuech, A. 2021. "Prise En Compte de La Qualité Des Données de Naissance Dans Le Calcul Des PAT Et Des Index IBOVAL." *Institut de l'Élevage*. <https://idele.fr/detail-article/prise-en-compte-de-la-qualite-des-donnees-de-naissance-dans-le-calcul-des-pat-et-des-index-iboval>.
- Kirkpatrick, S., C. D. Gelatt, and M. P. Vecchi. 1983. "Optimization by Simulated Annealing." *Science* 220 (4598): 671–80. <https://doi.org/10.1126/science.220.4598.671>.
- Xiang, Yang, Sylvain Gubian, Brian Suomela, and Julia Hoeng. 2013. "Generalized Simulated Annealing for Global Optimization: The GenSA Package." *The R Journal* 5 (1): 13–28.