



HAL
open science

Utilisation de l'imagerie 3D pour estimer la note d'état corporel des vaches laitières

Amélie Fischer

► **To cite this version:**

Amélie Fischer. Utilisation de l'imagerie 3D pour estimer la note d'état corporel des vaches laitières. Sciences agricoles. 2013. dumas-00940939

HAL Id: dumas-00940939

<https://dumas.ccsd.cnrs.fr/dumas-00940939>

Submitted on 5 Feb 2014

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.



AGROCAMPUS OUEST

CFR Rennes*

65, rue de Saint-Brieuc
35 000 RENNES



INRA

UMR PEGASE

Domaine de la
Prise
35590 Saint-Gilles

Mémoire de Fin d'Études

**Diplôme d'Ingénieur de l'Institut Supérieur des Sciences Agronomiques,
Agroalimentaires, Horticoles et du Paysage ***

Année universitaire* : 2012-2013

Spécialité* : SAED (Sciences animales pour l'élevage de demain)

Spécialisation ou option* : professionnel

**Utilisation de l'imagerie 3D pour estimer la note d'état corporel
des vaches laitières**

Par : Amélie FISCHER

Volet à renseigner par l'enseignant responsable de l'option/spécialisation
ou son représentant*

Date :/.../... Signature

Bon pour dépôt (version définitive)

Autorisation de diffusion : Oui Non

Devant le jury :

Soutenu à Rennes le 17/09/2013

Sous la présidence de* : Lucile MONTAGNE

Maître de stage* : Philippe FAVERDIN

Enseignant référent : Sandrine LAGARRIGUE

Autres membres du jury (Nom, Qualité) : Florence GARCIA, rapporteur

*"Les analyses et les conclusions de ce travail d'étudiant n'engagent
que la responsabilité de son auteur et non celle d'AGROCAMPUS OUEST".*

Fiche de confidentialité et de diffusion du mémoire

Cadre lié à la confidentialité :

Aucune confidentialité ne sera prise en compte si la durée n'en est pas précisée.

Préciser les limites de la confidentialité ⁽²⁾ :

➤ Confidentialité absolue : oui non

(ni consultation, ni prêt)

Si oui 1 an 5 ans 10 ans

Le maître de stage ⁽⁴⁾,

Cadre lié à la diffusion du mémoire :

A l'issue de la période de confidentialité et/ou si le mémoire est validé diffusable sur la page de couverture, il sera diffusé sur les bases de données documentaires nationales et internationales selon les règles définies ci-dessous :

Diffusion de la version numérique **du mémoire** : oui non

Référence bibliographique diffusable ⁽³⁾ : oui non

Résumé diffusable : oui non

Mémoire papier consultable sur place : oui non

Reproduction autorisée du mémoire : oui non

Prêt autorisé du mémoire papier : oui non

.....
Diffusion de la version numérique **du résumé** : oui non

↳ Si oui, l'auteur ⁽¹⁾ complète l'autorisation suivante :

Je soussigné(e) _____, propriétaire des droits de reproduction dudit résumé, autorise toutes les sources bibliographiques à le signaler et le publier.

Date : _____ Signature : _____

Rennes/Angers, le

L'auteur ⁽¹⁾,

L'enseignant référent
ou son représentant

(1) auteur = étudiant qui réalise son mémoire de fin d'études

(2) L'administration, les enseignants et les différents services de documentation d'AGROCAMPUS OUEST s'engagent à respecter cette confidentialité.

(3) La référence bibliographique (= Nom de l'auteur, titre du mémoire, année de soutenance, diplôme, spécialité et spécialisation/Option)) sera signalée dans les bases de données documentaires sans le résumé.

(4) Signature et cachet de l'organisme

REMERCIEMENTS

Je remercie Philippe FAVERDIN pour m'avoir accueillie au sein de l'UMR PEGASE et m'avoir soutenu tout au long du stage. Merci pour son accompagnement attentif dans toutes les étapes du stage et pour ses conseils pertinents et son aide pendant ces 6 mois.

Je tiens sincèrement à remercier 3DOuest pour avoir rendu la réalisation du stage possible et m'avoir accueillie une journée à Lannion : merci à Jean-Michel DELOUARD pour son accueil chaleureux, à Laurent DELATTRE et Thibault Luginbühl pour avoir été disponibles à chaque instant et patients pour m'initier à l'imagerie 3D.

Je remercie également toute l'équipe de la station expérimentale de Méjusseume pour leur disponibilité permanente et pour leur bonne humeur : merci à Jacques LASSALAS pour son accueil , à Michel FARGETTON qui a toujours répondu présent pour m'aider au bon fonctionnement des mesures, Ginette THEAUD qui quel que soit le niveau d'anticipation de ma demande m'a accueillie au sein du système de gestion des données de l'installation, à Arnaud MOTTIN pour avoir rendu possible les manipulations sur les vaches, à Jérémy ESLAN pour la conception du système d'acquisition et à André COZIEN, Mathieu GUILLOUX et Patrick PICHOT pour avoir partagé avec moi les secrets de la notation d'état corporel des vaches.

Je remercie aussi Erwan CUTULLIC pour ses conseils précieux dans la rédaction d'un script avec le logiciel R.

Je remercie Sandrine LAGARRIGUE pour son suivi du bon fonctionnement du stage.

Merci à toute l'équipe de l'UMR PEGASE pour son accueil et sa bonne humeur.

Merci à ma famille pour leur soutien permanent et leur motivation dans les rares moments difficiles !

Enfin, je suis particulièrement reconnaissante envers mes colocataires, Hélène et Louise, pour tous les moments (sérieux et surtout les moins sérieux) partagés ensemble et qui m'ont supportée pendant les 6 mois !

Merci à vous tous !

SOMMAIRE

Glossaire

Liste des abréviations

Liste des figures

Liste des tableaux

| | |
|---|----|
| Introduction..... | 1 |
| Matériel et méthodes | 4 |
| 1. Le dispositif expérimental..... | 4 |
| a. La population étudiée | 4 |
| b. Les mesures de l'état corporel des vaches par des notateurs..... | 4 |
| c. Le système d'acquisition | 4 |
| 2. Le traitement des surfaces en 3 dimensions (3D) : des acquisitions brutes aux données utilisées pour l'estimation de la NEC | 5 |
| a. Sélection des surfaces 3D sur la qualité d'image..... | 5 |
| b. Normalisation des surfaces 3D | 5 |
| i) Changement de repère en 3D | 5 |
| ii) Alignement des surfaces | 6 |
| c. Recherche du repère optimal | 7 |
| 3. La construction de l'indicateur nec3d..... | 7 |
| a. La calibration de l'indicateur | 8 |
| i) Dispositif expérimental pour calibrer la nec3d | 8 |
| ii) Une calibration par régression linéaire multiple..... | 8 |
| b. Le processus de validation de la méthode | 9 |
| i) La population d'étude | 9 |
| ii) Validation de la méthode par analyse de l'erreur de prédiction du modèle | 9 |
| c. La qualification de la méthode..... | 10 |
| i) Dispositif expérimental pour évaluer la fidélité de la nec3d..... | 10 |
| ii) Définition statistique de la répétabilité et de la reproductibilité | 10 |
| d. Comparaison de différentes méthodes de pointage et de découpage du masque..... | 11 |
| Résultats..... | 11 |
| 1. Un prototype fonctionnel sans problème majeur..... | 11 |

| | | |
|----|---|----|
| 2. | Comparaison des 2 méthodes de pointage | 12 |
| a. | Des masques différents, mais de qualité de calibration très voisines..... | 12 |
| b. | Une sélection du meilleur modèle par stepwise différente de l'ordre des composantes principales | 12 |
| c. | Une validation externe de la nec3D nettement moins bonne que la calibration..... | 13 |
| 3. | Une méthode plus répétable et reproductible que la NEC | 14 |
| 4. | La résolution du masque a peu d'impact sur la valeur prédictive de la méthode « nec3D ». | 15 |
| | Discussion | 16 |
| 1. | L'influence du choix du paramétrage sur la nec3d..... | 17 |
| 2. | Une méthode répétable et reproductible | 18 |
| 3. | L'imagerie 3D : une technique avec des perspectives d'avenir | 19 |
| | Conclusion | 20 |
| | Références Bibliographiques | 21 |
| | Annexes | 25 |

GLOSSAIRE

AIC : Akaike Information Criterion : le meilleur modèle est celui avec l'AIC le plus bas. L'AIC est pénalisé par le nombre de variables dans le modèle, il ne diminue donc pas naturellement quand le nombre de variables dans le modèle augmente.

Biais ou erreur quadratique moyenne (MSE) : adéquation entre les valeurs théoriques et les valeurs mesurées (Wallach et al., 2006).

Discontinuité : une surface qui n'est pas lisse, c'est-à-dire contenant des ruptures. Elles apparaissent quand la vache a bougé brusquement lors de l'acquisition.

Gold standard : test qui fait référence dans un domaine pour établir la validité d'un fait.

Plan moyen : le plan dont la distance aux 4 points identifiés est minimale.

Régression des moindres rectangles : La régression basée sur la technique des moindres rectangles ajuste l'équation de régression aux données en minimisant la distance euclidienne de chacune des données à la droite de régression, au lieu de minimiser l'écart entre les valeurs observées et les valeurs prédites [3]

Répétabilité : qualité de l'accord entre des mesures d'un même échantillon effectuées dans des conditions expérimentales très faiblement variables. Elle s'obtient en répétant un protocole de mesure sur un même échantillon, l'appareillage et l'opérateur étant identiques et les mesures étant effectuées dans un faible intervalle de temps. Evidemment la répétabilité n'intègre pas les erreurs systématiques ou les défauts d'appareillage. Elle donne simplement une mesure de la dispersion du signal pour un protocole de mesure effectué sur un appareil donné et par un opérateur donné [2].

Reproductibilité : même définition que la répétabilité, mais dans le cas de mesures effectuées dans des conditions fortement variables au sein du même laboratoire [2].

Ridge régression ou régression pseudo-orthogonalisée : cette régression impose une longueur maximale à la norme du vecteur des coefficients du modèle de régression. En faisant varier cette longueur maximale, la régression permet d'identifier les variables instables, dont les coefficients sont instables et tendent vers zéro et dont les coefficients sont faibles (Tomassone et al., 1983).

Technique du « one-shot » : technique d'imagerie basée sur une unique prise de vue de l'animal et non sur la superposition de 90 prises de vue.

Zone cible : zone comprise entre les pointes de hanche et les pointes de fesses.

Zone d'ombre ou trou : une partie vide de la surface qui correspond à une zone qui n'est pas dans le champ de vision du capteur. Ces zones sont dépourvues de coordonnées, on peut qualifier ces acquisitions d'incomplètes.

LISTE DES ABRÉVIATIONS

3D : 3 dimensions

ACP : analyse en composantes principales

ANOVA : analyse de variance

CV_T / σ_r : coefficient de variation de répétabilité / écart-type de répétabilité

CV_R / σ_R : coefficient de variation de reproductibilité / écart-type de reproductibilité

etr : écart-type résiduel

NEC : note d'état corporel

Nec3d : estimateur de la note d'état corporel à partir des surfaces 3D

Repère « ACP » : repère issu de l'ACP

Repère « animal » : repère 3 dimensions centré sur l'animal

Repère « capteur » : repère 3 dimensions du capteur

RMSE : root mean square error

« 3points » : méthode de pointage basée sur 4 points dont 2 identifient la base du sacrum

« 4points » : méthode de pointage basée sur 4 points dont 2 identifient les pointes de fesses

« validdiff » : population de validation contenant des acquisitions de vaches non incluses dans la population de calibration

« valididem » : population de validation contenant des acquisitions de vaches de la population de calibration mais à un stade de lactation différent

LISTE DES FIGURES

Figure 1 : Angles mesurés à partir de l'image du dos de la vache, d'après Bewley et al. (2008)

Figure 2 : Modélisation polynomiale de la partie dorsale postérieure d'une vache maigre (gauche) et grasse (à droite), d'après Halachmi et al. (2008)

Figure 3 : Localisation de la surface (STI) et de l'angle (ATI) utilisés comme indicateurs (d'après Negretti et al., 2008)

Figure 4 : localisation des zones de palpation pour déterminer la note d'état corporel d'après les grilles RNED (1984)

Figure 5 : Localisation de la zone cible scannée par le capteur

Figure 6 : Qualité des acquisitions: surface lisse et complète (A), surface discontinue (B), surface incomplète (C)

Figure 7 : Localisation des 4 points identifiés pour les deux méthodes: « 4points » et « 3points »

Figure 8 : Identification du repère "animal"

Figure 9 : Homogénéisation du gabarit des vaches par l'étape de normalisation

Figure10 : Construction du masque contenant les points valorisés pour la calibration

Figure 11 : Description de la population de calibration : NEC, poids vif, rang de lactation

Figure 12 : Description des populations de validation, valididem (à gauche) et validdiff (à droite): NEC, poids vif et rang de lactation

Figure 13 : Description de la population "reproductibilité": NEC, poids moyen, rang de lactation

Figure 14 : Masques obtenus pour les deux méthodes de pointage: "4points" et "3points"

Figure 15 : Dimensions des 2 masques

Figure 16 : Evolution de l'inertie expliquée par les composantes principales de l'ACP pour les méthodes "4points" et "3points"

Figure 17 : Evolution de la qualité des modèles au cours des étapes de sélection « pas à pas » pour les 2 méthodes de pointage

Figure 18 : Qualité du modèle de calibration pour la méthode "4points" et la méthode "3points"

Figure 19 : Evolution des erreurs résiduelles de prédiction au cours des étapes de la sélection "pas à pas" pour les 2 populations de validation et les 2 méthodes de pointage

Figure 20 : Comparaison des résidus de prédiction de la nec3d des 2 méthodes de pointage pour les 2 populations de validation

Figure 21 : Comparaison des résolutions et du nombre de pixels renseignés pour les masques obtenus pour 3 résolutions: 150x150, 100x100, 50x50

LISTE DES TABLEAUX

Tableau 1: Comparaison de grilles de notation d'état corporel en fonction de leur pays d'utilisation, de l'échelle de notation, des zones anatomiques évaluées et du type de notation

Tableau 2: Comparaison des travaux réalisés sur l'estimation de la NEC par imagerie sur la nature des acquisitions, le pointage, l'échelle de notation, l'estimateur utilisé, la validation externe et la qualification de la méthode

Tableau 3: coordonnées des 4 points identifiés

Tableau 4: Organisation du plan de validation : exemple pour un niveau de NEC

Tableau 5: ordre d'apparition des composantes principales dans le modèle sélectionné "pas à pas" pour les deux méthodes de pointage

Tableau 6: Décomposition de l'erreur quadratique moyenne de l'estimation de la NEC pour les 2 méthodes de pointage et les 2 populations de validation externe

Tableau 7: Résultats de validation des modèles associés aux 3 résolutions: 150x150, 100x100 et 50x50 pour les 2 populations de validation, valididem et validdiff

Tableau 8: Corrélation entre les résidus issus des prédictions réalisées par les modèles associés aux résolutions (150x150, 100x100 et 50x50) pour les 2 populations de validation (valididem et validdiff)

Tableau 9: Comparaison de la répétabilité et de la reproductibilité de la nec3d pour les résolutions de masque: 150x150, 100x100 et 50x50

Tableau 10: Comparaison de la répétabilité et de la reproductibilité des mesures de NEC par les notateurs et des estimations de NEC par la nec3d

Tableau 1: Comparaison de grilles de notation d'état corporel en fonction de leur pays d'utilisation, de l'échelle de notation, des zones anatomiques évaluées et du type de notation

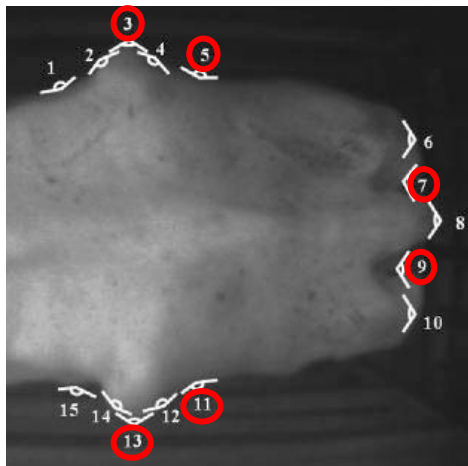
| Grille | Pays | Echelle | Notation |
|------------------------------|----------------------|---------|----------|
| Earle (1976) | Australie | 1 à 8 | V |
| Wildman et al. (1982) | Irlande | 1 à 5 | P ou V |
| ITEB (1984) | France | 0 à 5 | P ou V |
| Edmonson et al. (1989) | Etats-Unis | 1 à 5 | V |
| MacDonald et Roche (2004) | Nouvelle- Zélande | 1 à 10 | P |

V : méthode visuelle / P : méthode par palpation

INTRODUCTION

L'état des réserves corporelles des vaches laitières qui concerne les quantités de réserves énergétiques stockées dans le tissu adipeux, musculaire ou intermusculaire, est souvent un indicateur de l'état de l'animal en relation avec des critères zootechniques importants tels que l'état de santé ou les performances de reproduction de l'animal. En effet, des vaches maigres ou grasses ont des performances de reproduction plus faibles que des vaches en bon état corporel : leur taux de réussite en 1^{ère} insémination artificielle est plus faible, leur intervalle vêlage-vêlage est plus long et le retour en chaleur est plus précoce (Pryce et al., 2000 ; Berry et al., 2003 ; Dechow et al., 2002). Ruegg et Milton (1995) ont observé que les vaches malades ont en moyenne une note d'état corporel inférieure de 0,25 point à celle des vaches saines. La force de l'association entre un état corporel et l'état de santé des vaches est à nuancer avec le type de maladie observée (Gillund et al., 2001 ; Roche et Berry, 2006). La sélection génétique des vaches laitières Prim'Holstein a augmenté la valeur génétique du troupeau laitier en termes de production, mais a détérioré les critères de reproduction et de santé. Ces associations entre caractères zootechniques et état des réserves corporelles motivent la prise en compte de l'état des réserves corporelles dans le calcul de l'indice de sélection pour améliorer les performances de reproduction et de santé des vaches tout en conservant un niveau de production élevé (Berry et al., 2003 ; Coffey et al., 2003 ; Pryce et Harris, 2006).

L'appréciation de l'état des réserves corporelles est cependant difficile. Une multitude de méthode de mesure d'état des réserves corporelles existe, soit par mesure directe passant par la composition de la carcasse, soit par mesure indirecte. L'estimation directe n'est pas applicable à grande échelle et est lourde et coûteuse étant donné qu'elle implique l'abattage des animaux et la dissection anatomique de la carcasse (Robelin et al., 1982 ; Szabo et al., 1999). Néanmoins c'est la méthode de mesure la plus précise, elle peut servir de méthode de référence pour calibrer d'autres méthodes. Parmi les autres méthodes, certaines sont précises mais chères, longues et invasives pour l'animal et l'observateur, telles que l'estimation par le diamètre des adipocytes ou la technique de l'eau lourde qui rend la viande impropre à la consommation (Waltner et al., 1994). D'autres sont précises, répétables et faciles à utiliser, mais impliquent d'avoir la technologie nécessaire pour réaliser les mesures, c'est le cas pour les mesures de gras sous-cutané par échographie (Schröder et Staufenbiel, 2006). Actuellement, en élevage, les mesures d'état corporel sont effectuées avec la note d'état corporel (NEC) parce qu'elle est simple, pratique, non invasive et la moins chère de toutes les méthodes. Chez la vache laitière, son estimation est appréciée par des techniciens soit par palpation de zones anatomiques prédéfinies, soit visuellement. Une NEC faible décrit une vache maigre alors qu'une note élevée décrit une vache très grasse. Différentes grilles de notation d'état corporel ont été mises en place, elles divergent au niveau de l'échelle de notation et de la précision des points de la note à attribuer (tableau 1). En France, le réseau des chambres d'agriculture et l'INRA utilisent la grille définie par Bazin (1984). Cette grille décrit une note allant de 0 à 5 avec la possibilité d'attribuer des demi-points pour chacune des zones anatomiques. La note finale est la moyenne des notes attribuées pour chacune des zones (annexe I). Cependant la subjectivité liée à l'interprétation des grilles d'évaluation par les



- Angles retenus comme indicateur
 - 1-15 : angle antérieur de la pointe de la hanche
 - 2-14 : courbure antérieure pointe de la hanche
 - 3-13 : angle de la pointe de la hanche
 - 4-12 : courbure postérieure de la pointe de la hanche
 - 5-11 : angle postérieur de la pointe de la hanche gauche
 - 6-10 : angle de la pointe des fesses
 - 7-9 : détroit caudal
 - 8 : angle de la base de la queue
- Avec a-b : les angles respectivement associés au côté droit et gauche

Figure 1: Angles mesurés à partir de l'image du dos de la vache, d'après Bewley et al. (2008)

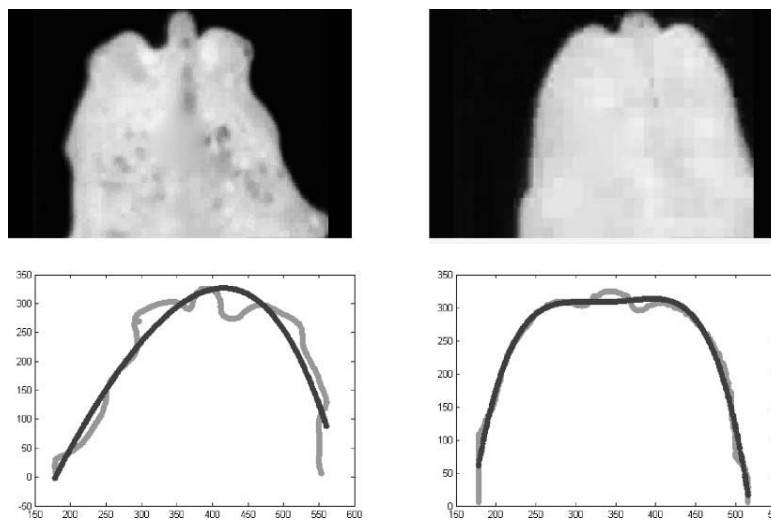


Figure 2: Modélisation polynomiale de la partie dorsale postérieure d'une vache maigre (gauche) et grasse (à droite), d'après Halachmi et al. (2008)

Les images du haut sont brutes, acquises par la caméra thermique et celles du bas sont les graphiques issus de la modélisation polynomiale : contour du dos de la vache et le polynôme.

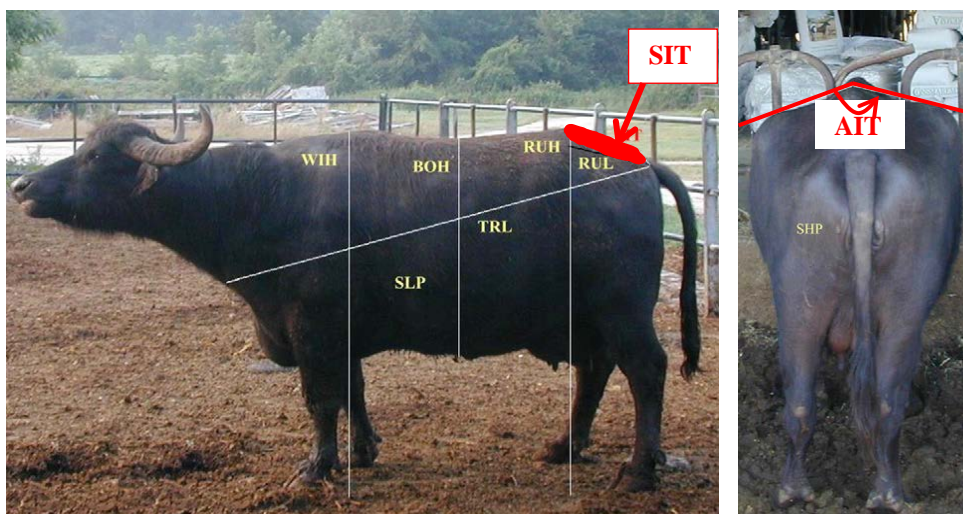


Figure 3: Localisation de la surface et de l'angle utilisés comme indicateurs (d'après Negretti et al., 2008)

SIT : surface de la tubérosité iliaque /

AIT : angle entre le sommet de la croupe et les demi-tangentes aux tubérosités iliaques

techniciens ainsi que la nécessité de ré-étalonner les techniciens pour s'assurer de la qualité de leur notation et son manque de précision sont des critères à améliorer pour utiliser la NEC comme méthode de référence. A cet effet, de récents travaux ont étudié la possibilité d'automatiser les mesures de note d'état corporel.

Dans un objectif d'automatisation et d'objectivation de la notation de l'état corporel, plusieurs études ont été menées pour estimer l'état corporel des vaches en utilisant des techniques d'imagerie. Ces techniques sont de plus en plus abordables financièrement et ont une qualité d'image et une précision qui en font un outil potentiellement intéressant. Différentes approches dont l'objectif est d'estimer l'état corporel des vaches à partir d'images sont référencées dans la littérature : l'image peut être prise sous forme de photo ou bien d'image infra-rouge. L'analyse peut se concentrer sur l'ensemble du dos de l'animal (Azzaro et al., 2011) ou bien sur la partie arrière de la vache qui constituent les zones les plus souvent utilisées pour les méthodes de notation d'état corporel. Le traitement des images va de la simple notation visuelle des photos jusqu'à des traitements sophistiqués. Les notes d'état attribuées à partir de photos reflètent fidèlement les notes d'état obtenues *de visu*. Ferguson et al. (2006) concluent que la notation d'état corporel peut se faire à partir de photographie de la partie dorsale postérieure de la vache. Néanmoins, la subjectivité de la notation persiste et l'image n'apporte pas un avantage déterminant. Les travaux postérieurs ont valorisé l'image en 2 dimensions (2D) des vaches comme un support d'informations pour estimer la NEC des vaches. La difficulté de cette approche consiste dans un premier temps à définir le système d'acquisition utilisé : un système à bas coût basé sur des photographies peut être sensible aux variations environnementales. Bewley et al. (2008) ont par exemple été confronté à un appareillage sensible aux variations de luminosité et n'ont pu conserver que 23% de leurs images. Halachmi et al. (2008) affirment que les résultats obtenus avec un nombre limité d'images ne reflètent pas les capacités maximales de la méthode. L'estimation de la NEC passe par la définition d'un ensemble de caractéristiques issues des images 2D qui sont estimées comme potentiellement corrélées à la note d'état corporel. En ce sens, Bewley et al. (2008), Halachmi et al. (2008) et Negretti et al. (2008) ont défini des critères concrets relatifs à la forme du dos de la vache : Bewley et al. (2008) ont testé l'estimation de la NEC par des angles localisés autour de la pointe de hanche, de l'attache de la queue et de la pointe des fesses ; Halachmi et al. (2008) ont testé l'estimation de la NEC par l'écartement du contour du dos de la vache à la parabole modélisant au mieux ce contour et Negretti et al. (2008) ont testé l'estimation de la NEC par des surfaces (figures 1 à 3). Bewley et al. (2008) ont observé que 3 angles étaient fortement corrélés à la NEC : l'angle de la pointe de hanche ($r = 0,48$) ; l'angle postérieur à la pointe de hanche ($r = 0,52$) et l'angle du détroit caudal ($r = 0,31$; figure 1). Ces angles sont plus obtus pour des vaches grasses et plus aigus pour des vaches maigres. La note d'état corporel a été estimée à partir des 2 angles de hanches avec un modèle ajustant 92,79% des notes de calibration avec une erreur de 0,25 point. De la même manière, Negretti et al. (2008) ont obtenu une très bonne calibration entre la NEC mesuré par les techniciens et la NEC estimée à partir de l'angle et de la surface de la tubérosité iliaque ($R^2 = 0,85$ et $rmse = 0,3$). Halachmi et al. (2008) n'ont obtenu qu'un coefficient de corrélation de 0,315 entre la NEC mesurée et estimée par leur indicateur. Parmi ces 3 travaux, seul Negretti et al. (2008) a réalisé une validation externe pour évaluer la qualité de prédiction de la NEC sur une

Tableau 2: Comparaison des travaux réalisés sur l'estimation de la NEC par imagerie sur la nature des acquisitions, le pointage, l'échelle de notation, l'estimateur utilisé, la validation externe et la qualification de la méthode

| | Nature des acquisitions | Pointage (oui/non) | Echelle de notation | Estimateur de NEC | Validation externe | Qualification de la méthode |
|------------------------|-------------------------|----------------------------|--|---|--------------------|-----------------------------|
| Ferguson et al. (2006) | Image en 2D | Non | 1 à 5 (Ferguson et al., 1994) | NEC visuelle à partir des photos du dos de la vache | Non | Non |
| Bewley et al. (2008) | Image en 2D | Oui (23 points identifiés) | 0 à 5 (Lowman et al., 1976) 1 à 5 (Edmonson et al., 1989) | Angles | Non | Non |
| Halachmi et al. (2008) | Image infra-rouge en 2D | Oui (23 points identifiés) | 1 à 5 (Edmonson et al., 1989) | Ecart du contour du dos de la vache au polynôme modélisant au mieux ce contour | Non | Non |
| Negretti et al. (2008) | Image en 2D | Non | 1 à 9 (Campanile et al., 1998) | Surfaces de zones anatomiques Angles | Oui | Non |
| Azzaro et al. (2011) | Image en 2D | Oui (23 points identifiés) | 1 à 5 (Ferguson et al., 1994) | Ecart de la forme du contour du dos de la vache à la forme moyenne du contour dans la population de calibration | Oui | Non |
| Projet NEC3D (2013) | Surfaces 3D | / | / | NEC3D | Oui | Oui |

population différente de la population de calibration : 72% des NEC de la population de validation sont prédites avec une erreur inférieure à 4,5%. Azzaro et al. (2011) ont poursuivi les recherches pour améliorer l'estimation de la NEC à partir d'informations issues d'images, en valorisant davantage toute l'information disponible dans les images. Ils exploitent les images non pas à travers une grandeur unique comme des angles ou des surfaces mais à travers la position des 23 points anatomiques définis par Bewley et al. (2008) (figure 1). La variabilité des NEC mesurées est traduite en variabilité de la forme du contour du dos de la vache, c'est-à-dire en variabilité de la position de ces 23 points. Ils utilisent une analyse en composante principale (ACP) pour calculer la forme moyenne du contour de la population de calibration. L'ACP est couramment utilisé en analyse de forme en imagerie (Cootes et al., 1992 ; Luginbühl, 2012). C'est l'écart d'une nouvelle forme à cette forme moyenne qui estime la NEC. Ils comparent l'estimation faite par l'indicateur testé par Bewley et al. (2008), celui testé par Halachmi et al. (2008) avec celle faite par leurs indicateurs : un dont les données sont traitées par une **ACP à noyau polynomial**¹ et un dont les données sont traitées par une ACP. L'estimation réalisée par l'indicateur relatif au traitement par une ACP à noyau est celle qui prédit la NEC avec l'erreur la plus faible en validation par **LOOCV**². Des progrès ont été réalisés en matière d'automatisation de mesure de NEC, mais une étape demeure manuelle : l'identification de points anatomiques pour normaliser les images (tableau 2).

L'objectif du projet consiste à mettre au point, en collaboration étroite avec la société 3Douest, une nouvelle méthode de notation de l'état corporel en utilisant les nouvelles possibilités de l'imagerie en 3 dimensions (3D). L'information contenue dans des images 3D est plus complète que celle contenue dans les images 2D et offre donc des possibilités intéressantes pour évaluer le relief de la partie arrière des vaches qui est la zone la plus utilisée pour la notation de l'état corporel. Au-delà de la définition du cahier des charges de ce dispositif et de son implémentation dans les installations expérimentales de Méjusseume, le travail de ce stage vise 1/ à tester la possibilité de calibrer les notes d'état corporel des vaches à partir de l'image 3D, 2/ à affiner la méthode du traitement de ces images (points de repère, résolution finale de l'image, équations de calibration) et 3/ à qualifier cette nouvelle méthode en termes de répétabilité, de reproductibilité et de qualité de prédiction des NEC.

1 Une ACP à noyau polynomial est une ACP appliquée sur des données qui ont au préalable été transformée par une fonction polynomiale (Hastie et al., 2009).

2 LOOCV : leave one out cross validation : la validation est réalisée sur une seule observation de la population et la calibration sur les autres observations. Ce procédé, aussi appelé Jackknife, est répété jusqu'à ce que chacune des observations de la population ait été utilisée comme observation de validation.

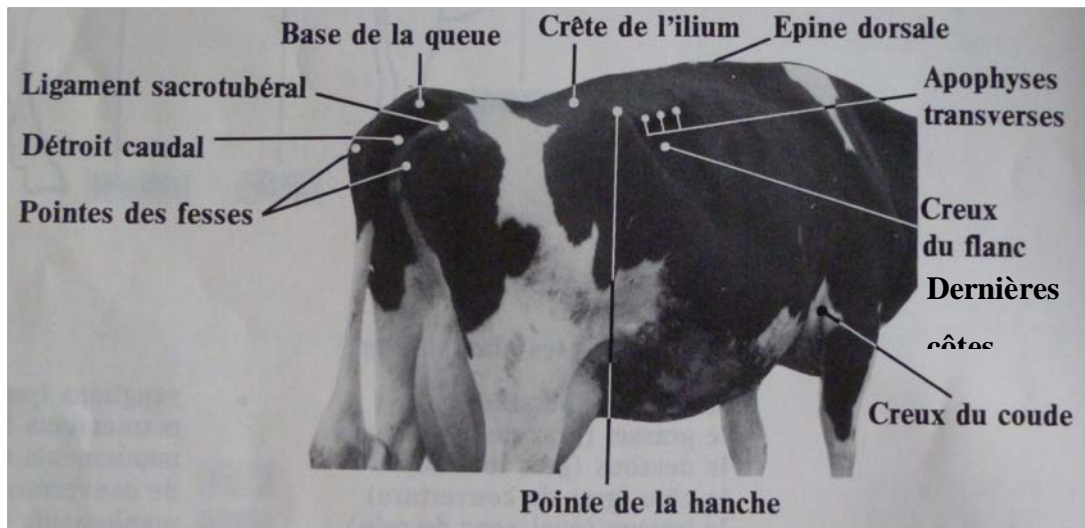


Figure 4: Localisation des zones de palpation pour déterminer la note d'état corporel d'après les grilles de Bazin (1984)

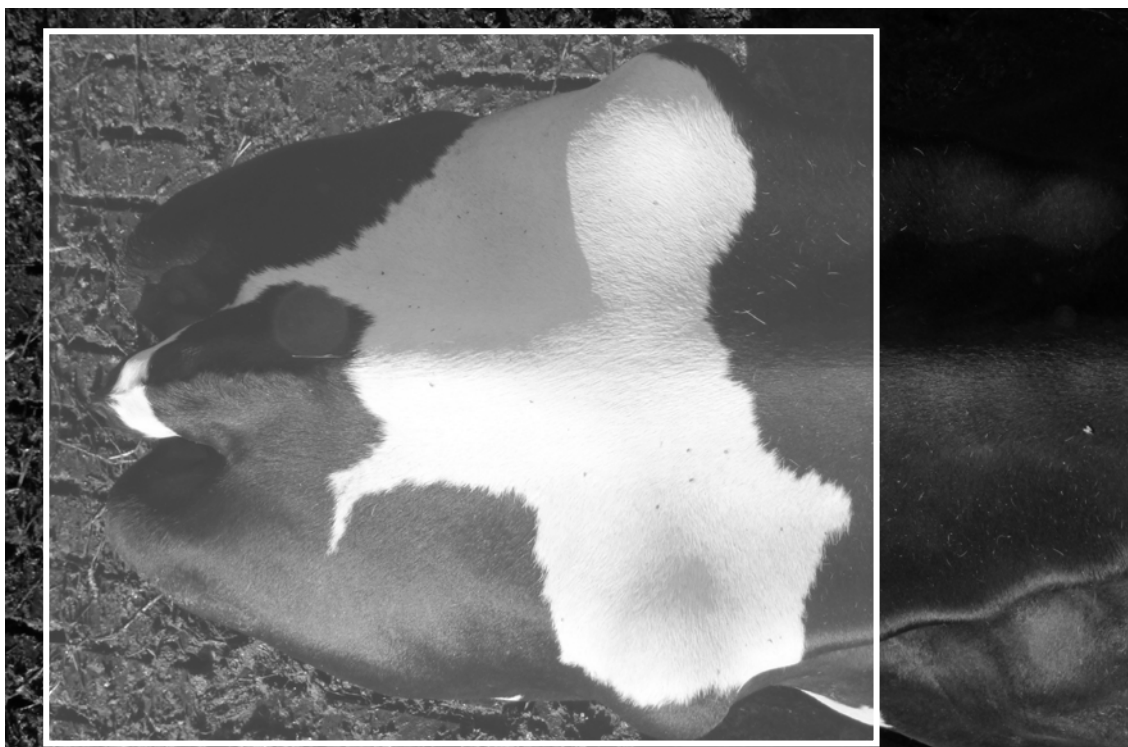


Figure 5: Localisation de la zone cible scannée par le capteur

MATÉRIEL ET MÉTHODES

1. Le dispositif expérimental

a. La population étudiée

Les mesures ont été réalisées sur le troupeau de vaches laitières de l'installation expérimentale INRA de Méjusseume. Au total, le troupeau compte environ 180 vaches de race Prim'Holstein avec un renouvellement moyen de 40%, produisant en moyenne environ 8500 kg de lait par vache et par an. Les vaches sont traitées deux fois par jour dans une salle de traite rotative avec traite par l'arrière (De Laval). A chaque traite, les vaches sont pesées et identifiées automatiquement sur un poste de pesée électronique.

b. Les mesures de l'état corporel des vaches par des notateurs

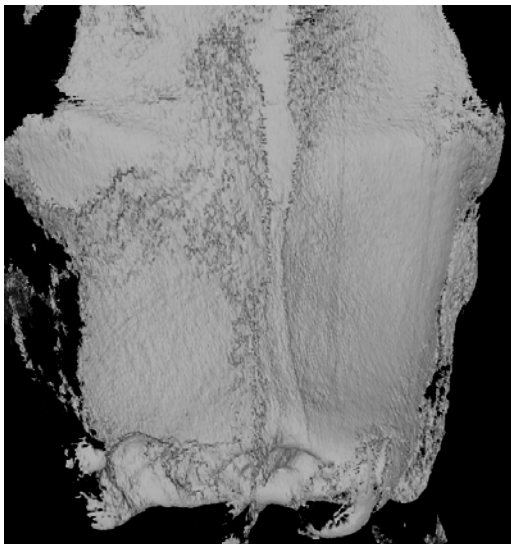
L'état corporel de toutes les vaches du troupeau est noté une fois par mois par 3 techniciens, selon les grilles définies par l'Institut technique de l'élevage bovin (Bazin, 1984). A l'installation expérimentale de Méjusseume, les notes sont attribuées par palpation de 2 zones anatomiques : les dernières côtes et le ligament sacro-tubéral (figure 4). Les notes s'échelonnent de 0 à 5 avec la possibilité d'attribuer des quarts de point par zone. La note d'état corporel obtenue est la moyenne des notes par zone analysée et la NEC est la moyenne des NEC attribuées par les 3 techniciens.

c. Le système d'acquisition

Le système d'acquisition et de traitement des images a été développé par la société 3D Ouest (J.M. Delouard, L. Delattre, T. Luginbühl) suivant un cahier des charges établi en collaboration avec l'INRA pour répondre à l'objectif de notation d'état corporel des vaches. Les surfaces 3D ont été réalisées par un capteur Xtion PRO Live Motion Sensor (ASUSTeK Computer Inc.) placé sur la balance à la sortie du roto. De cette manière, les vaches sont immobilisées pour la pesée, ce qui permet de réaliser des acquisitions 3D de meilleure qualité. Le contrôle du capteur se fait à distance depuis un ordinateur. L'acquisition démarre automatiquement à la fermeture des portes de la balance, par l'intermédiaire d'un capteur mécanique placé sur la porte d'entrée de la balance (annexe II).

La zone cible retenue pour estimer la NEC des vaches laitières est centrée sur la partie dorsale du bassin. La zone scannée par le capteur est une zone élargie par rapport à la zone cible allant au-delà des pointes de hanche et en deçà des pointes de fesses pour être sûr de capturer la zone cible pour chaque vache (figure 5). En focalisant le capteur sur cette zone anatomique, l'idée est de valoriser une zone commune aux grilles de NEC. De plus, cette zone dorsale postérieure est plus facilement accessible et moins sensible aux mouvements des vaches dans la stalle de pesée.

Le capteur est fixé à 2,02 m du sol indépendamment du squelette de la balance, en aval de la porte d'entrée. La zone cible est définie dans les 3 dimensions et n'est pas totalement accessible d'un seul point de vue. Le capteur est par conséquent muni d'un mouvement rotatif lorsqu'une vache rentre dans la balance, lui permet de scanner la partie arrière de la vache, des pointes de hanche vers les pointes de fesses à raison de 30 images par seconde. Les vaches sont identifiées manuellement depuis l'ordinateur en parallèle de l'acquisition par le biais d'une webcam placée au-dessus des vaches au milieu de la balance. Chaque nouvelle



2
trous

Figure 6 : Qualité des acquisitions: surface lisse et complète (A), surface discontinue (B), surface incomplète (C)

acquisition enregistrée contient la date, l'heure à la seconde près et le numéro d'identification de la vache.

Toutes les étapes allant du pointage des 4 points à l'estimation de la nec3d pour une population externe à la population de calibration à l'ACP sont réalisées par le programme conçu par T. Luginbühl (3DOuest). Toutes les analyses statistiques ont été réalisées avec le logiciel R (3.0.1, 2013).

2. Le traitement des surfaces en 3 dimensions (3D) : des acquisitions brutes aux données utilisées pour l'estimation de la NEC

a. Sélection des surfaces 3D sur la qualité d'image

Les surfaces 3D brutes, c'est-à-dire enregistrées après l'acquisition, sont visualisables avec le logiciel Meshlab v1.3.0.® et sont monochromes pour éviter que l'opérateur soit influencé par la robe de la vache lors du traitement (figure 6). Chaque fichier est visualisé afin de juger de la qualité des acquisitions : si la surface présente des **discontinuités**³ ou si elle possède des **zones d'ombre**⁴, c'est-à-dire qu'elle est incomplète, alors la surface est de mauvaise qualité (figure 6). Les zones d'ombre sont communes sur les surfaces 3D mais c'est leur taille et leur nombre qui définissent la qualité de l'acquisition. Les acquisitions utilisées dans les populations d'étude sont toutes caractérisées par des surfaces lisses et ont été sélectionnées en évitant des images avec des zones d'ombre trop importantes en nombre et en taille.

b. Normalisation des surfaces 3D

En analyse d'image, la normalisation est une étape importante qui permet de définir des points de référence identiques à toutes les images, c'est-à-dire de rendre les données comparables (Krukowski M., 2009). Les vaches n'ont pas la même position dans la balance, donc dans le repère du capteur. De plus, elles se distinguent par leur gabarit qui ne doit pas intervenir dans la notation de l'état corporel. Au final, les surfaces ne sont pas dans un repère centré sur l'animal et n'ont pas toutes la même taille. La normalisation va permettre de définir une zone commune à étudier sur chaque acquisition et de calibrer ces acquisitions sur une même métrique, pour que les données analysées par la suite soient indépendantes du gabarit de la vache (Azzaro et al., 2011). Cette normalisation passe dans un premier temps par un changement de repère caractérisé par le passage du repère du capteur, repère « capteur », à un repère « animal » centré sur la surface 3D et dans un second temps par une série de transformations pour ajuster les différents gabarits sur une même métrique. Un logiciel spécifiquement développé par T. Luginbühl pour ce travail a permis de réaliser cette normalisation des surfaces.

i) Changement de repère en 3D

Pour définir un nouveau repère, il faut définir au moins 2 axes du repère, le troisième étant orthogonal aux 2 premiers. Les 2 axes doivent être facilement identifiables, nous avons choisi

³ Une discontinuité caractérise ici une surface qui n'est pas lisse, c'est-à-dire contenant des ruptures. Elles apparaissent quand la vache a bougé brusquement lors de l'acquisition.

⁴ Une zone d'ombre ou trou définit une partie vide de la surface qui correspond à une zone qui n'est pas dans le champ de vision du capteur. Ces zones sont dépourvues de coordonnées, ces acquisitions sont incomplètes.

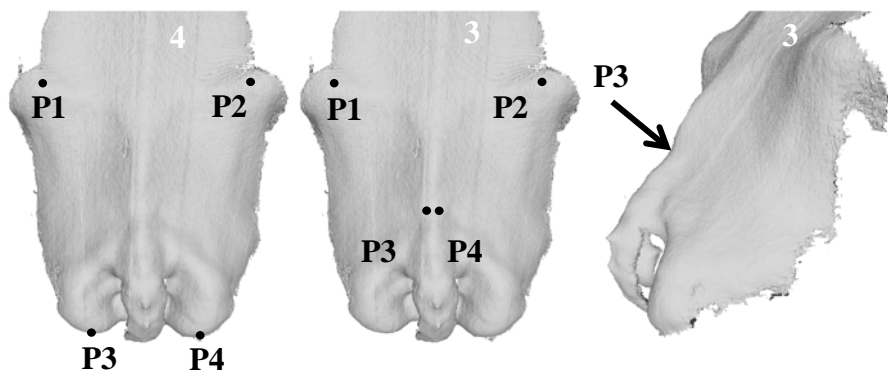


Figure 7: Localisation des 4 points identifiés pour les deux méthodes:
 « 4points » et « 3points »
 4 : méthode « 4points » / 3 : méthode « 3points »

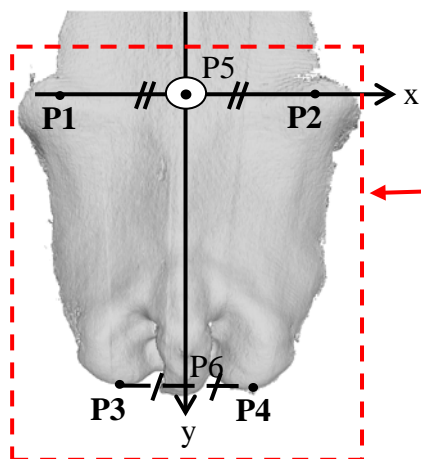


Figure 8: Identification du repère "animal"
 ⊙ Axe z vient vers le lecteur et est perpendiculaire aux axes x et y

Tableau 3: Coordonnées des 4 points identifiés

| | « 4points » | « 3points » |
|----|----------------|--------------|
| P1 | (0, -250, 0) | (0, -250, 0) |
| P2 | (0, 250, 0) | (0, 250, 0) |
| P3 | (500, -125, 0) | (280, -8, 0) |
| P4 | (500, 125, 0) | (280, 8, 0) |

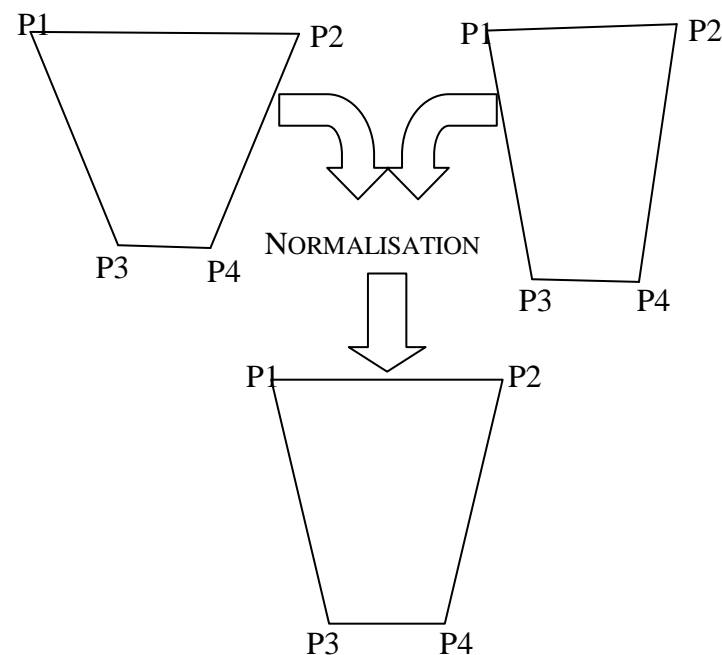


Figure 9 : Homogénéisation du gabarit des vaches par l'étape de normalisation

l'axe de la colonne vertébrale pour l'axe des X et l'axe passant par les 2 pointes de hanche pour l'axe des Y. En pratique, l'identification de 3 points suffit pour définir un plan. En se restreignant à 3 points, la définition du plan est plus sensible aux erreurs d'identification des 3 points, c'est pourquoi nous avons choisi d'identifier 4 points. Le plan formé par les axes X et Y du repère est le **plan moyen**⁵ défini par les 4 points. La difficulté a été de trouver 4 points (P1, P2, P3, P4) facilement repérables et identifiables sur les surfaces 3D. Nous avons testé 2 méthodes se distinguant par la position des points P3 et P4. Dans les 2 cas, P1 et P2 identifient respectivement la pointe de hanche gauche et la pointe de hanche droite. Selon la première méthode, notée « 4points », P3 et P4 identifient respectivement les pointes de fesses gauche et droite. Selon la deuxième méthode, P3 et P4 sont très proches, séparés d'environ 3cm, et se situent à la base du sacrum. Cette zone est caractérisée par une dépression, un « creux » local (figure 7). Cette méthode, parce que P3 et P4 sont très proches, est censée reproduire un système basé sur l'identification de 3 points, elle est notée « 3points ». La comparaison des 2 méthodes a pour objectif d'évaluer si la méthode « 3points » apporte une estimation de la NEC d'aussi bonne qualité que la méthode « 4points », autrement dit s'il est possible de supprimer l'identification manuelle d'un point pour gagner du temps et en répétabilité de pointage. L'axe X est défini par la projection du vecteur $\overrightarrow{P1P2}$ sur le plan moyen et l'axe Y est défini par la projection du vecteur $\overrightarrow{P5P6}$ sur le plan moyen avec P5 le milieu du segment P1P2 et P6 le milieu du segment P3P4. L'axe Z est orienté vers le haut (figure 8).

L'alignement de ces repères est réalisé par une série de rotations cherchant dans un premier temps à aligner l'axe X du repère « animal » avec l'axe X du repère « capteur » et dans un second temps à rendre le plan moyen orthogonal à l'axe Z du repère « capteur ». Les surfaces sont ensuite transformées linéairement avec la contrainte d'amener les points P1, P2, P3 et P4 à des coordonnées fixes et identiques pour toutes les surfaces. Cette transformation modifie les coordonnées relatives aux axes X et Y des points de la surface, mais conserve les coordonnées en Z. En pratique, après la transformation, les points P1, P2, P3 et P4 de toutes les surfaces se superposent et leurs coordonnées ont été choisies comme résultant de la moyenne de leurs coordonnées avant la transformation dans la population étudiée (tableau 3).

A l'issue de cette transformation, toutes les surfaces 3D à analyser ont le même gabarit. En effet, les dimensions du quadrilatère P1P2P3P4, assimilables au gabarit des vaches, sont identiques pour toutes les surfaces après la transformation linéaire (figure 9). Les différences existantes entre les différentes surfaces ne seront donc pas liées à une différence de gabarit, mais bien à des différences d'état corporel.

ii) Alignement des surfaces

La surface 3D conservée pour l'analyse, appelé le masque, est réduite à un rectangle focalisée sur une zone plus étendue que la **zone cible**⁶ (figure 8). Le nombre de points renseignés dans ce masque est un facteur modulable : le masque est découpé selon l'axe des X et l'axe des Y pour former une grille. Par défaut cette grille a une dimension de 150x150, elle renseigne

⁵ Le plan moyen est le plan dont la distance aux 4 points identifiés est minimale.

⁶ Zone cible : zone comprise entre les pointes de hanche et les pointes de fesses.

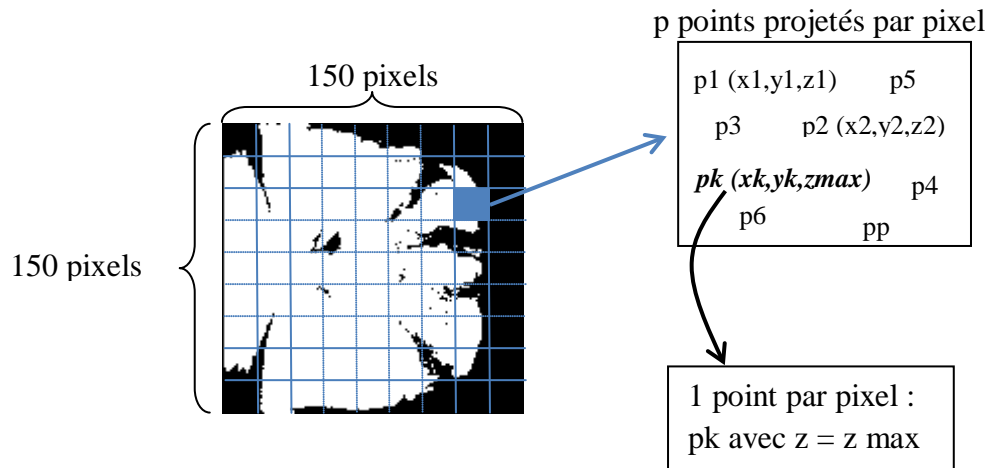


Figure 10 : Construction du masque contenant les points valorisés pour la calibration

22 500 points en tout. Chaque pixel de cette grille contient plusieurs points du masque et ne renseigne au final que les coordonnées du point possédant la coordonnée maximal pour Z (figure 10). Quand un pixel est localisé sur une zone d'ombre, il possède les coordonnées moyennes des 8 pixels voisins, le point renseigné dans le pixel est le barycentre des pixels voisins. Pour éviter que les données traitées ne contiennent de valeurs manquantes lorsqu'il y a des zones d'ombre importantes, seuls les points communs à toutes ces surfaces font partie du masque et sont utilisés pour les calculs.

c. Recherche du repère optimal

Après cet alignement, les surfaces normalisées sont projetées dans le repère « animal » qui est également un espace en 3D. Ces surfaces contiennent une information géométrique corrigée des différences de localisation, d'alignement et de gabarit. L'ensemble des surfaces utilisées pour construire l'indicateur (jeu de calibration) sont traitées par une ACP pour trouver la transformation qui mette le mieux en évidence les variations existantes entre les surfaces 3D et qui possède le moins de dimensions (Luginbühl, 2012). Le traitement des acquisitions par l'ACP est un procédé courant en imagerie 3D afin de réduire l'information des coordonnées des surfaces à celle des composantes principales (Cootes et al, 1992 ; Luginbühl, 2012).

Les surfaces 3D sont projetées dans le repère de l'ACP dont les composantes principales définissent les axes du repère et sont indépendantes les unes des autres. Au total il y a autant de composantes principales que de surfaces 3D incrémentées dans l'ACP. Chaque surface est donc définie par une combinaison linéaire de ses coordonnées dans ce nouveau repère, c'est cette information condensée qui va être utilisée pour prédire les NEC.

3. La construction de l'indicateur nec3d

La mise en place d'un indicateur passe par 3 phases : une phase de calibration, une phase de validation et une phase de qualification de la méthode (Feinberg M., 2010). La calibration consiste à qualifier le lien existant entre les données issues des surfaces 3D et leur NEC correspondantes. La validation permet de vérifier si la méthode, telle qu'elle est définie à l'issue de la calibration, fournit des estimations de NEC appelées « nec3d » en accord avec les NEC mesurées par les notateurs pour des acquisitions n'ayant pas servi à la calibration. La définition de la méthode s'accompagne de sa caractérisation, notamment en termes de fidélité, pour juger l'étroitesse de l'accord entre les nec3d obtenues pour des acquisitions répétées d'une même vache dans des conditions spécifiées (Feinberg M., 2010). La fidélité s'estime par la **répétabilité**⁷ et la **reproductibilité**⁸. Comme les mesures sont toutes réalisées sur le site de Méjusseume, nous ne mesurerons que la reproductibilité intra-installation.

⁷ Répétabilité : qualité de l'accord entre des mesures d'un même échantillon, effectuées dans des conditions expérimentales très faiblement variables. Elle s'obtient en répétant un protocole de mesure sur un même échantillon, l'appareillage et l'opérateur étant identiques et les mesures étant effectuées dans un faible intervalle de temps. Evidemment la répétabilité n'intègre pas les erreurs systématiques ou les défauts d'appareillage. Elle donne simplement une mesure de la dispersion du signal pour un protocole de mesure effectué sur un appareil donné et par un opérateur donné [2].

⁸ Reproductibilité : même définition que la répétabilité, mais dans le cas de mesures effectuées dans des conditions fortement variables au sein du même laboratoire [2].

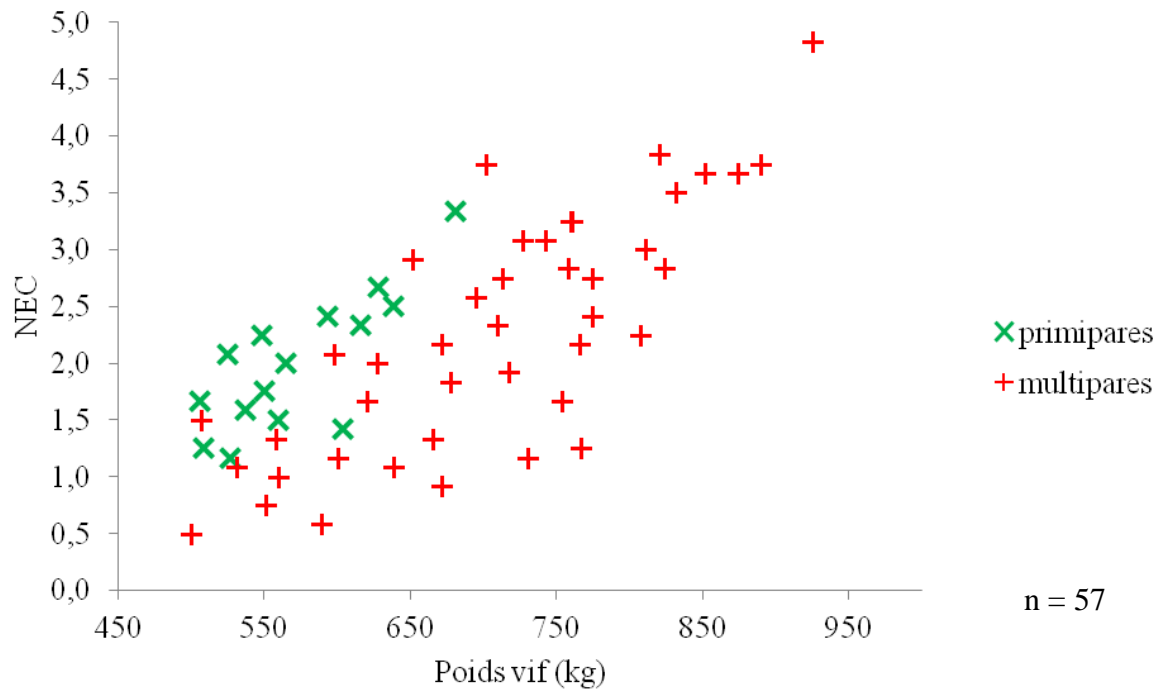


Figure 8: Description de la population de calibration : NEC, poids vif, rang de lactation

a. La calibration de l'indicateur

i) Dispositif expérimental pour calibrer la nec3d

Des acquisitions ont été réalisées mensuellement le jour de notation d'état corporel pour toutes les vaches en lactation lors de la traite du soir. Pour être sûr d'avoir au moins une acquisition de bonne qualité pour chaque vache et pour chaque série de notation, les acquisitions du jour de notation ont été complétées par celles du jour précédent ou suivant la notation. Ces séries d'acquisitions servent à la définition des populations d'acquisitions destinées d'une part à la calibration de l'indicateur, notée population « calibration » et d'autre part à la validation de l'indicateur.

Les vaches ont dans l'ordre été sélectionnées sur la NEC, le poids vif, le rang de lactation et le stade de lactation dans le but de définir une population hétérogène en termes d'âge et de format et possédant la gamme de NEC la plus large possible. Etant donné que les vaches extrêmes en termes de NEC, c'est-à-dire avec une NEC inférieure à 1 ou supérieure à 3, sont peu nombreuses, elles ont toutes été utilisées dans la population « calibration » pour que l'indicateur soit construit sur une gamme de NEC la plus large possible.

La population « calibration » contient 57 acquisitions relatives à 56 vaches différentes dont la NEC s'échelonne de 0,5 à 4,8 avec une moyenne à 2,2 (figure 11). Les acquisitions utilisées pour définir ces populations ont en priorité été sélectionnées sur la qualité des surfaces. Pour une même série de NEC, 2 acquisitions par vache sont disponibles : celle du jour de notation et celle du jour précédent ou suivant. C'est l'acquisition du jour de notation qui a été favorisée pour la sélection. Néanmoins si cette surface n'était pas de bonne qualité, c'est l'acquisition possédant le moins de trous et/ou avec la surface la plus lisse qui est conservée pour l'analyse.

ii) Une calibration par régression linéaire multiple

Les données de ces acquisitions constituent la matrice de l'ACP et définissent le repère utilisé. Chacune des 57 acquisitions est donc décrite par 57 coordonnées correspondant à sa projection sur les 57 composantes principales de l'ACP, appelé repère « ACP ». La NEC est estimée par régression linéaire multiple de ces coordonnées en utilisant la fonction « lm » du logiciel R. Le modèle complet de régression linéaire (1) cherche à expliquer les variations de la NEC mesurée par les notateurs à partir des coordonnées des surfaces 3D dans le repère « ACP » :

$$NEC_A = \alpha + \beta_{A1} \times a_{A1} + \beta_{A2} \times a_{A2} + \dots + \beta_{A57} \times a_{A57} + \varepsilon \quad (1)$$

Avec $(a_{A1}, a_{A2}, a_{A3}, \dots, a_{A57})$ le vecteur des 57 coordonnées de la surface 3D de la vache A.

NEC_A : la NEC de la vache A

ε , les résidus du modèle, α la constante du modèle et les β_{ij} , les coefficients du modèle pour la vache i et la coordonnée relative à l'axe j du repère.

Le modèle (1) est le modèle complet, contenant les 57 variables relatives aux coordonnées dans le repère « ACP ». Ce modèle complet peut cependant être surparamétré et peut intégrer des variables ayant peu d'intérêt pour prévoir la NEC. Le meilleur modèle est sélectionné par la méthode pas à pas (stepwise) avec la fonction « step » du logiciel R qui, comme la méthode pas à pas ascendante, part du modèle nul et introduit à chaque étape la variable explicative la plus significative au seuil de 5% lorsqu'on régresse la variable à expliquer (NEC) sur cette variable et toutes les variables explicatives sélectionnées dans les étapes précédentes (Lafaye

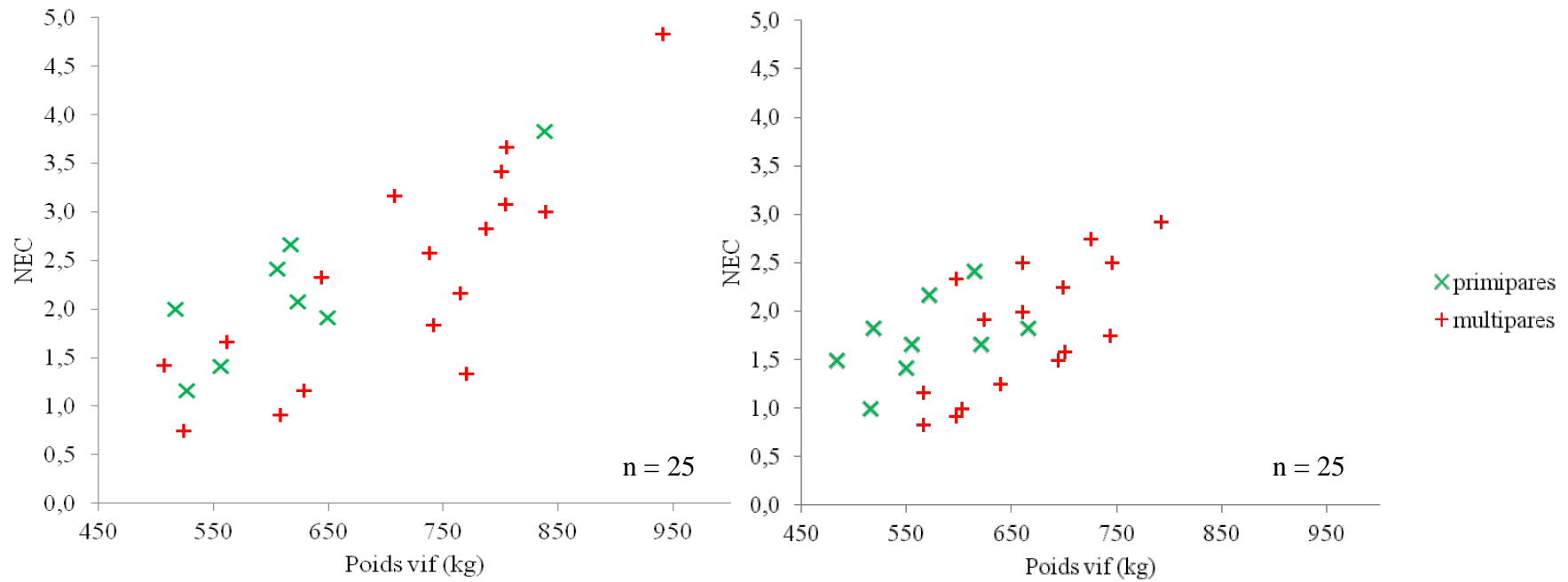


Figure 9: Description des populations de validation, valididem (à gauche) et validdiff (à droite): NEC, poids vif et rang de lactation

de Micheaux et al., 2012). La régression stepwise, contrairement à la méthode pas à pas ascendante, ne sélectionne que des variables significatives au seuil de 5% et élimine des variables déjà intégrées dans le modèle dans les étapes précédentes si ces variables sont moins informatives que la dernière variable sélectionnée. La sélection s'arrête quand le critère de sélection des variables, en l'occurrence l'AIC⁹ (Akaike Information Criterion), n'évolue plus. Le modèle issu de chaque étape du stepwise est également testé sur les populations de validation et de répétabilité. L'évolution à chaque étape du stepwise des coefficients de détermination ajusté et des erreurs de prédiction sur ces populations permet d'évaluer la qualité prédictive du modèle. Si le modèle est surparamétré, le modèle prédisant la NEC sur une population externe à celle de calibration avec l'erreur de prédiction la plus faible possède moins de composantes principales que le modèle sélectionné par stepwise. Le modèle retenu à l'issue de cette sélection est le modèle « stepwise ».

b. Le processus de validation de la méthode

i) La population d'étude

Les acquisitions utilisées dans le procédé de validation sont des acquisitions qui n'ont pas servi à calibrer l'indicateur nec3d. A cet effet, 2 populations de validation sont constituées : l'une constituée d'acquisitions de vaches de la population de calibration, mais prises à un stade de lactation différent, c'est la population « valididem » et l'autre possédant des acquisitions de vaches différentes de la population « calibration », c'est la population « validdiff ». Chacune des populations de validation contient 25 acquisitions relatives à 25 vaches. La population « valididem » possède une gamme allant de 0,8 à 4,8 avec une moyenne de 2,3 alors que la population « validdiff » est caractérisée par une gamme de NEC allant de 0,8 à 2,9 avec une moyenne de 1,8, les vaches avec des notes extrêmes ayant été utilisées pour le jeu de calibration (figure 12). Au final, ces populations ont été définies à partir de 4 séries de NEC relatives aux mois d'avril, mai, juin et juillet. La différence de stade de lactation des vaches incluses dans la population de calibration et dans la population « valididem » varie de 1 à 4 mois.

ii) Validation de la méthode par analyse de l'erreur de prédiction du modèle

La validation de la méthode consiste à évaluer la capacité du modèle sélectionné en phase de calibration à faire des prédictions de bonne qualité pour une population externe à la population de calibration. En pratique, la validation repose sur l'adéquation entre les nec3d calculés par le modèle sélectionné en calibration et les NEC mesurées par les notateurs. La qualité de l'adéquation est évaluée avec le **biais**¹⁰, l'erreur de prédiction et le coefficient de détermination (R^2) de la régression de la NEC en fonction de la nec3d (Wallach et al., 2006). L'erreur moyenne de prédiction de la NEC, EMP, est la racine carrée du biais inexpliqué, B_i . Ces différents critères caractérisent le modèle de régression linéaire qui explique les variations de la NEC à partir des variations des valeurs estimées, nec3d (5) :

⁹ AIC : Akaike Information Criterion : le meilleur modèle est celui avec l'AIC le plus bas. Il est pénalisé par le nombre de variables dans le modèle, il ne diminue donc pas naturellement quand le nombre de variables dans le modèle augmente.

¹⁰ Biais ou erreur quadratique moyenne (MSE) mesure l'adéquation entre les valeurs théoriques et les valeurs mesurées (Wallach et al., 2006).

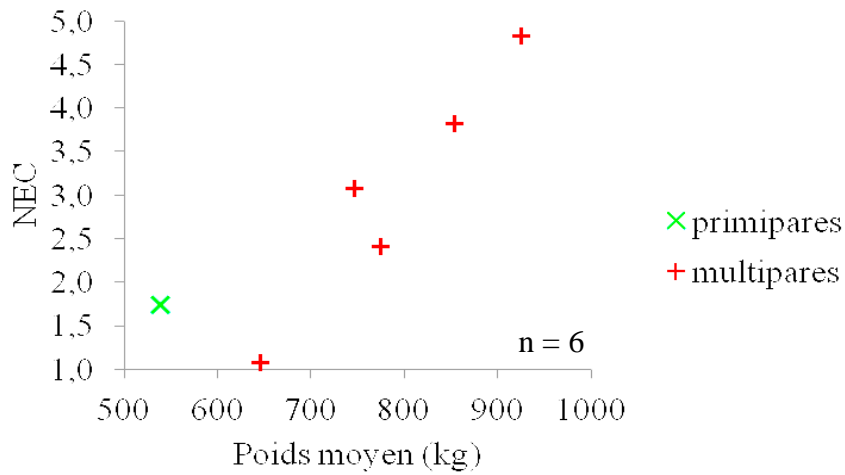


Figure 10: Description de la population "reproductibilité": NEC, poids moyen, rang de lactation

Le poids moyen est celui des 8 séries d'acquisition pour chaque vache.

Tableau 4: Organisation du plan de validation : exemple pour un niveau de NEC

| NEC | Série | Mesurage (nec3d) | | | | | | | |
|------|-------|------------------|----------|----------|----------|----------|----------|----------|----------|
| | | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 |
| 1,08 | 1 | x | x | x | x | x | x | x | x |
| | 2 | x | | | | | | | |
| | 3 | x | | | | | | | |
| | 4 | x | | | | | | | |
| | 5 | x | | | | | | | |
| | 6 | x | | | | | | | |
| | 7 | x | | | | | | | |
| | 8 | x | | | | | | | |

Chaque série est caractérisée par le passage des 6 vaches de la population.

Texte : données utilisées pour l'analyse de la répétabilité.

Texte : données utilisées pour l'analyse de reproductibilité.

$$NEC = \alpha + \beta \times nec3d + \varepsilon \quad (5)$$

Avec β la pente de la régression, α l'ordonnée à l'origine et ε les résidus du modèle.

Le biais ou MSE est défini par la somme de 3 composantes : le **biais de la moyenne**, noté B_m , défini par la différence moyenne entre les valeurs mesurées et les valeurs calculées ; le **biais de la pente**, noté B_p , défini par l'écart à 1 de la pente de la régression de la NEC en fonction de la nec3d ; le **biais inexpliqué**, noté B_i , qui caractérise la corrélation existante entre les variations de la NEC et les variations de la nec3d.

c. La qualification de la méthode

La qualification de la méthode d'estimation de la NEC n'a été réalisée que pour la méthode « 3points ». L'objectif est d'évaluer l'erreur de mesure de la nec3d indépendamment de l'erreur de prédiction liée à l'erreur de la méthode de référence, la NEC.

i) Dispositif expérimental pour évaluer la fidélité de la nec3d

Des acquisitions répétées d'un même groupe de vaches ont été réalisées pour tester la répétabilité et la reproductibilité intra-installation de l'estimation de la NEC par la nec3d. Pour constituer le jeu de surfaces de ce test, 6 vaches ont été scannées 10 fois d'affilée, sans pause entre deux séries consécutives. Ces 6 vaches représentent des NEC différentes espacées régulièrement et décrivant une large gamme allant de 1,1 à 4,8. Pour une même NEC, les vaches sélectionnées sont selon les disponibilités caractérisées par des poids vifs différents et des rangs de lactation différents. Cette population est appelée population « reproductibilité » dans la suite du rapport (figure 13). Au final, seules les séries d'acquisition possédant des surfaces 3D de bonne qualité pour les 6 vaches sont utilisées. Les acquisitions conservées pour définir cette population sont toutes issues des mêmes séries d'acquisition. L'analyse porte donc sur 8 séries d'acquisitions de 6 vaches.

ii) Définition statistique de la répétabilité et de la reproductibilité

La répétabilité consiste à évaluer la capacité de la méthode à prédire la même nec3d pour une même acquisition : chacune des 6 vaches de la population de répétabilité est décrite par une seule acquisition de la population « reproductibilité » dont le pointage est répété 8 fois par le même opérateur (tableau 4). L'évaluation de la répétabilité de la nec3d teste donc la répétabilité de l'identification des 4 points. Nous cherchons à savoir dans quelle mesure nous sommes capables à partir de différents pointages consécutifs d'une même surface de prédire la même NEC. Au total, il y a 48 acquisitions dans la population, c'est la population « répétabilité ».

La reproductibilité intra-installation se traduit ici par la capacité de la méthode à prédire la même nec3d à partir de différentes acquisitions de la même vache réalisées sur une courte période de temps et pointées par le même observateur. Chacune des séries est donc à la fois différente des autres par le pointage effectué et par la surface 3D utilisée (tableau 4). La reproductibilité contient donc l'erreur de répétabilité à travers l'effet du pointage. L'écart entre la reproductibilité et la répétabilité renseigne la capacité du système d'acquisition à fournir la même information pour une même vache possédant la même NEC : plus l'écart est grand, moins le système d'acquisition est reproductible.

La répétabilité et la reproductibilité s'expriment par les écarts-type de répétabilité, noté σ_r , et

de reproductibilité intra-installation, noté σ_R . En retirant l'effet « vache » de la variation de la nec3d dans la population « répétabilité », l'effet du pointage est défini dans les résidus du modèle d'analyse de variance (ANOVA) à un facteur, l'effet vache. Ce facteur a 6 niveaux, correspondant aux 6 vaches de la population. σ_r est l'écart-type des résidus de ce modèle. σ_R est calculée de la même manière, en réalisant la même ANOVA sur la population « reproductibilité ».

La répétabilité et la reproductibilité s'évaluent respectivement avec le coefficient de variation de répétabilité, CV_r et de reproductibilité, CV_R :

$$CV_r = \frac{\sigma_r}{M_r} \times 100 \qquad CV_R = \frac{\sigma_R}{M_R} \times 100$$

avec M_r et M_R respectivement la nec3d moyenne dans la population « répétabilité » et « reproductibilité ».

La méthode sera d'autant plus répétable et reproductible que leur coefficient de variation est petit. La répétabilité σ_r caractérise la répétabilité globale du pointage des acquisitions. Les ANOVA sont réalisées avec la fonction « aov() » du logiciel R.

d. Comparaison de différentes méthodes de pointage et de découpage du masque

Les résultats de validation servent de base pour la comparaison des 2 méthodes de pointage et des 3 résolutions du masque : 150x150, 100x100, 50x50. Le modèle « stepwise » est ajusté pour chacune des 2 méthodes de pointage et pour chacune des 2 populations de validation, « validdiff » et « valididem ». Au total, l'étape de validation étudie 4 modèles de validation. Une régression linéaire des résidus de la prédiction effectuée par le modèle « stepwise » relatifs à une méthode de pointage en fonction des résidus relatifs à l'autre méthode de pointage est réalisée pour chacune des populations de validation afin de voir si ces erreurs sont indépendantes ou non. Cette régression linéaire ajuste le modèle selon la technique des **moindres rectangles**¹¹ car les deux méthodes jouent un rôle symétrique [3]. Le coefficient de corrélation associé à ces modèles sera d'autant plus élevé que le lien entre les résidus sera fort. La même méthode est appliquée pour conclure sur la possibilité de se restreindre à un masque renseignant 2 500 points à la place de 22 500 points.

La fonction « lm » complétée par le calcul de la pente et de l'ordonnée à l'origine d'après la technique des moindres rectangles et « cor » du logiciel R ont été utilisées respectivement pour la régression et les coefficients de corrélation. Toutes les analyses statistiques réalisées avec le logiciel R sont renseignées dans le script complet (annexe III).

RÉSULTATS

1. Un prototype fonctionnel sans problème majeur

Toutes les séries d'acquisition ont été réalisées avec le même prototype. Une acquisition dure 3 secondes et demande 1 seconde supplémentaire pour l'enregistrement sur l'ordinateur. Une à deux vaches par série d'acquisition n'ont pas été scannées à cause de panne temporaire du

¹¹ La régression basée sur la technique des moindres rectangles ajuste l'équation de régression aux données en minimisant la distance euclidienne de chacune des données à la droite de régression, au lieu de minimiser l'écart entre les valeurs observées et les valeurs prédites [3]

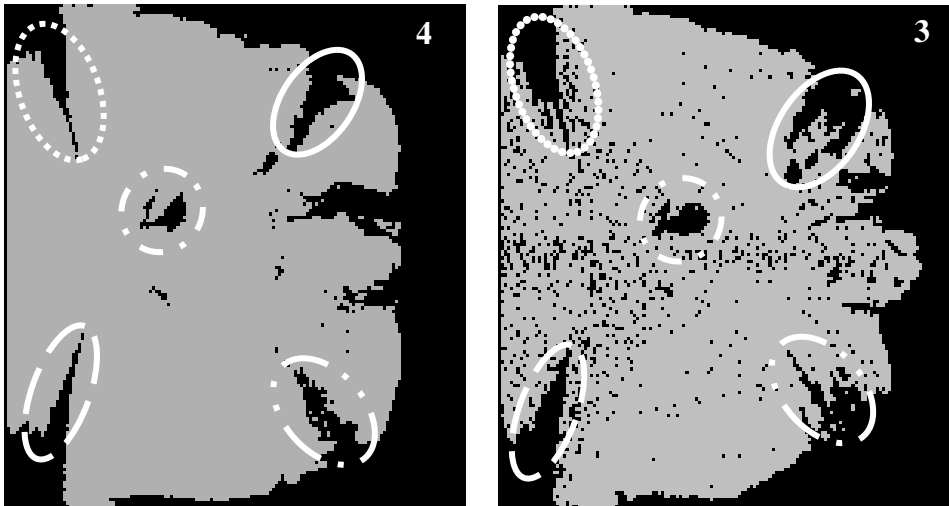


Figure 14 : Masques obtenus pour les deux méthodes de pointage: "4points" et "3points"
 4 : masque obtenu pour la méthode « 4points » 3 : masque obtenu pour la méthode « 3points »
 ○ Localisation de zones d'ombre du capteur. Chaque localisation du masque 4 a son équivalent sur le masque 3, représenté avec le même type d'ellipse.

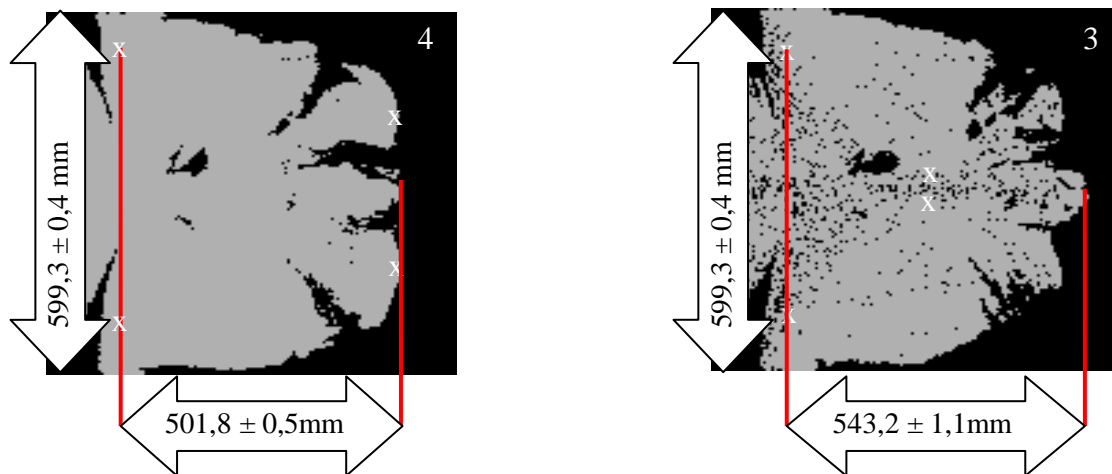


Figure 15 : Dimensions des 2 masques
 x : position des 4 points identifiés
 4 : méthode « 4points » / 3 : méthode « 3points »

programme d'acquisition. Ces pannes étaient rares et ne duraient que le temps équivalent à la réalisation d'une acquisition afin de le rendre à nouveau fonctionnel.

En 4 mois, 24 séries d'acquisitions ont été réalisées sur le troupeau en lactation de l'installation expérimentale de Méjusseume, c'est-à-dire 18 semaines consécutives dans la lactation comptabilisant un total de 3593 acquisitions comptabilisant en moyenne 10% de surfaces discontinues. Ces problèmes de discontinuités des surfaces 3D relèvent de mouvements trop brusques de la vache durant l'acquisition. Lorsque les vaches étaient stressées, notamment quand leur environnement était modifié (alimentation, logement, premières traites), elles bougeaient beaucoup durant l'acquisition. Vu la rapidité de mesure et la fréquence possible de ces acquisitions après automatisation, ce taux d'échec ne pose pas de problème majeur pour son développement. Une détection automatique des images difficilement exploitables serait sans doute nécessaire à l'avenir.

2. Comparaison des 2 méthodes de pointage

L'analyse de surfaces 3D passe par une phase de normalisation qui permet de définir une zone commune à toutes les surfaces analysées pour traiter la même information sur chacune des surfaces. La définition de cette zone commune, le **masque**, passe par l'identification de 4 points caractéristiques. La méthode « 4points » identifie les pointes de hanche et les pointes de fesses et la méthode « 3points » identifie les pointes de hanche et la base du sacrum. La comparaison des 2 méthodes de pointage montre malgré de petites différences des qualités assez voisines sur la calibration sur la NEC et la qualité prédictive des modèles « stepwise ».

a. Des masques différents, mais de qualité de calibration très voisines

La superposition des surfaces de la population de calibration aboutit à 2 masques différents pour les 2 méthodes (figure 14). Le masque pour le pointage « 3points » est parsemé de « trous », il est moins complet que le masque de la méthode « 4points ». Les surfaces analysées dans la méthode « 3points » contiennent donc moins d'information que la méthode « 4points » : chacune des surfaces 3D utilisée pour la calibration renseigne 13 847 points contre 15 231 points pour la méthode « 4points ». La superposition des 2 masques met en évidence des zones d'ombre communes témoignant de zones d'ombre du champ de vision du capteur (figure 14). Ces zones d'ombre sont présentes sur les surfaces quelle que soit leur qualité en termes de discontinuité. Cette différence de quantité d'informations contenue dans les surfaces 3D est complétée par une différence de dimensions du masque. Les 2 masques ont la même largeur au niveau des pointes de hanche, mais diffèrent au niveau de la longueur du bassin (figure 15). Le masque du pointage « 4points » est plus court que le masque du pointage « 3points », par conséquent les surfaces issues du pointage « 4points » sont plus comprimées que les surfaces issues du pointage « 3points ».

b. Une sélection du meilleur modèle par stepwise différente de l'ordre des composantes principales

Les ACP réalisées sur les populations de calibration pour chacune des méthodes définissent un nouveau repère à 57 axes, représentant les 57 composantes principales. Ce repère est celui qui décrit au mieux la variabilité existante dans la population de surfaces 3D de calibration en termes de forme. La régression linéaire multiple ajustée sur cette population aboutit à un modèle expliquant au mieux la variabilité observée de la NEC dans cette population.

Tableau 5: ordre d'apparition des composantes principales dans le modèle sélectionné "pas à pas" pour les deux méthodes de pointage

| méthode | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 |
|---------|-----|-----|-----|-----|-----|-----|-------------|-------------|-------------|-------------|
| 4points | cp1 | cp2 | cp5 | cp8 | cp7 | cp6 | cp10 | cp3 | <i>cp27</i> | <i>cp41</i> |
| 3points | cp1 | cp2 | cp6 | cp5 | cp9 | cp7 | <i>cp29</i> | <i>cp15</i> | <i>cp11</i> | <i>cp14</i> |

cpX : composante principale numéro X / *Texte* : composante principale ne faisant pas partie des 10 premières composantes de l'ACP

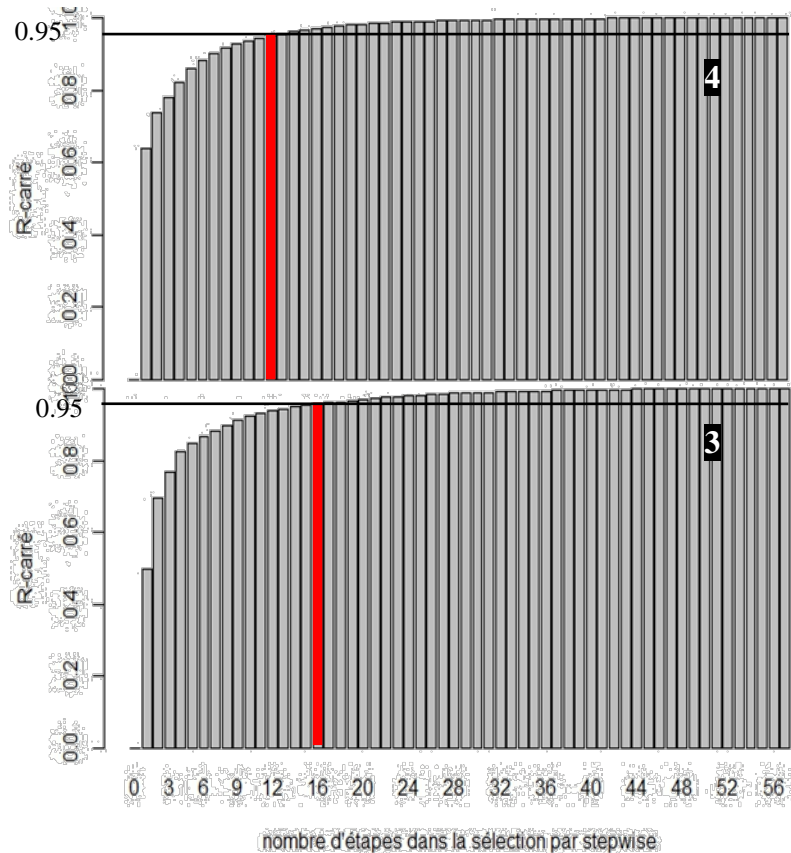


Figure 16 : Evolution de la qualité des modèles au cours des étapes de sélection « pas à pas » pour les 2 méthodes de pointage
 4 : méthode « 4points » / 3 : méthode « 3points »
 ■ Identification du modèle expliquant 95% de la variabilité observée de la NEC

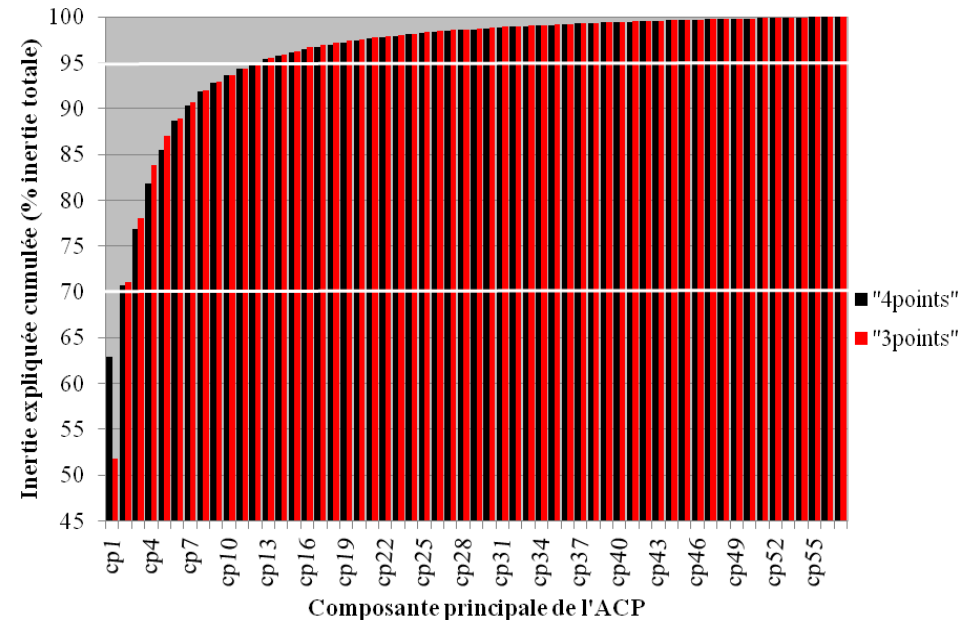


Figure 17 : Evolution de l'inertie expliquée par les composantes principales de l'ACP pour les méthodes "4points" et "3points"

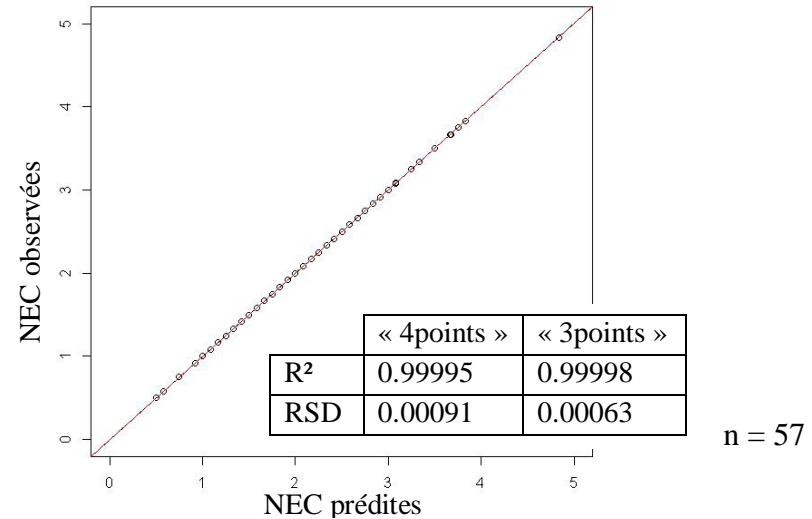


Figure 18 : qualité du modèle de calibration pour la méthode "4points" et la méthode "3points"

Chaque composante principale de l'ACP est caractérisée par l'inertie qu'elle explique, c'est-à-dire la part de variabilité existante entre les surfaces 3D de la population de calibration. Les composantes principales sont rangées dans l'ordre décroissant de l'inertie qu'elles expliquent. La comparaison de l'ordre des composantes principales dans l'ACP avec l'ordre d'apparition dans le modèle issu de la sélection « pas à pas » montre que seules les 2 premières composantes de l'ACP figurent dans le même ordre et au même rang dans le modèle, pour les 2 méthodes de pointage. La variabilité des formes 3D expliquée par ces 2 composantes, respectivement pour la méthode « 4points » et « 3points » est de 72,8% et 68,5% et la variabilité de la NEC expliquée par ce modèle à 2 composantes est de 71% (figure 16). L'ordre d'apparition des différentes composantes principales de l'ACP dans le modèle de la sélection « pas à pas » diffère de l'ordre des composantes principales de l'ACP à partir de la deuxième étape de la sélection. Parmi les 10 premières composantes ajoutées dans le modèle par la sélection figure les premières composantes principales de l'ACP, mais aussi des composantes principales survenant après la dixième composante principale : dans les 10 premières composantes du modèle, 4 composantes pour la méthode « 3points » et 2 pour la méthode « 4points » ne font pas partie des 10 premières composantes de l'ACP (tableau 5). Au final, 95% de l'inertie de la forme des surfaces 3D est expliquée par les 13 premières composantes de l'ACP alors que les modèles de régression contiennent respectivement pour la méthode « 4points » et « 3points » 12 et 16 composantes dans un ordre différent de celui de l'ACP pour expliquer 95% de l'inertie de la NEC (figure 16 & 17). Les composantes rajoutées par la suite n'expliquent qu'une petite part de cette inertie.

La régression linéaire multiple complétée de la sélection « pas à pas » du meilleur modèle aboutit à un modèle contenant 55 composantes principales pour les 2 méthodes de pointage. La sélection du meilleur modèle basée sur l'optimisation de l'AIC conserve le maximum de composantes principales qu'elle peut mettre dans le modèle sous la contrainte des 56 degrés de liberté du modèle. Au final, le modèle conservé pour l'estimation de la NEC est celui contenant 55 composantes principales excluant la 53^{ème} et la 16^{ème} pour la méthode « 4points » et la 57^{ème} et la 38^{ème} pour la méthode « 3points ». L'ajustement obtenu entre les NEC observées et les NEC prédites par les modèles de calibration est de très bonne qualité avec un coefficient de détermination proche de 1 et une erreur résiduelle (etr) proche de 0 (figure 18).

c. Une validation externe de la nec3D nettement moins bonne que la calibration

L'objectif de la validation externe est d'évaluer la qualité prédictive de la nec3D pour une population externe à la population de calibration. La première étape est d'analyser l'évolution de la qualité prédictive des modèles créés à chaque étape de la sélection « pas à pas » sur les 2 populations de validation afin de déterminer si le modèle conservé à l'issue de la calibration correspond au meilleur modèle en termes de qualité prédictive. Les erreurs résiduelles des modèles de prédiction diminuent curvilinéairement avec le nombre de composantes dans le modèle. Quelle que soit la méthode analysée, l'évolution de la qualité prédictive des modèles du stepwise indique une erreur résiduelle minimale pour des modèles contenant respectivement pour la méthode « 4points » et « 3points » 34 et 32 composantes pour la population « validdiff » et 10 et 39 composantes pour la population « valididem » (figure 19). Le meilleur modèle optimisé sur la qualité de prédiction sur les populations de validation

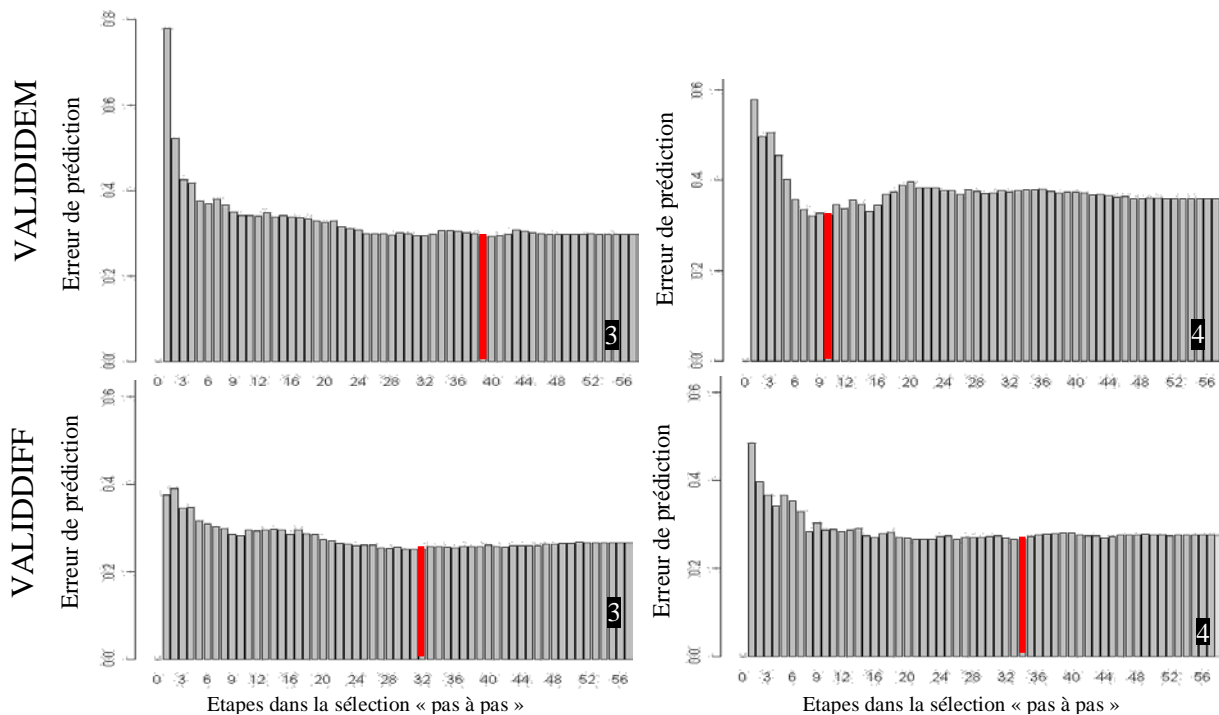


Figure 19 : Evolution des erreurs résiduelles de prédiction au cours des étapes de la sélection "pas à pas" pour les 2 populations de validation et les 2 méthodes de pointage
 4 : méthode « 4points » / 3 : méthode « 3points » / ■ erreur de prédiction minimale

Tableau 6: Décomposition de l'erreur quadratique moyenne de l'estimation de la NEC pour les 2 méthodes de pointage et les 2 populations de validation externe

| population | méthode | MSE | Bm (% MSE) | Bp (% MSE) | Bi (% MSE) | R ² ajusté | erreur (RMSE) |
|------------|---------|--------|------------|------------|------------|-----------------------|---------------|
| valididem | 4points | 0,1349 | 0,0402 | 0,0004 | 0,9594 | 0,8735 | 0,3672 |
| | 3points | 0,0935 | 0,0007 | 0,0449 | 0,9544 | 0,9127 | 0,3058 |
| validdiff | 4points | 0,1123 | 0,0335 | 0,2928 | 0,6737 | 0,7813 | 0,3351 |
| | 3points | 0,1004 | 0,0082 | 0,2883 | 0,7036 | 0,7959 | 0,3168 |

MSE : erreur quadratique moyenne / RMSE : root mean squared error : $\sqrt{\text{MSE}}$

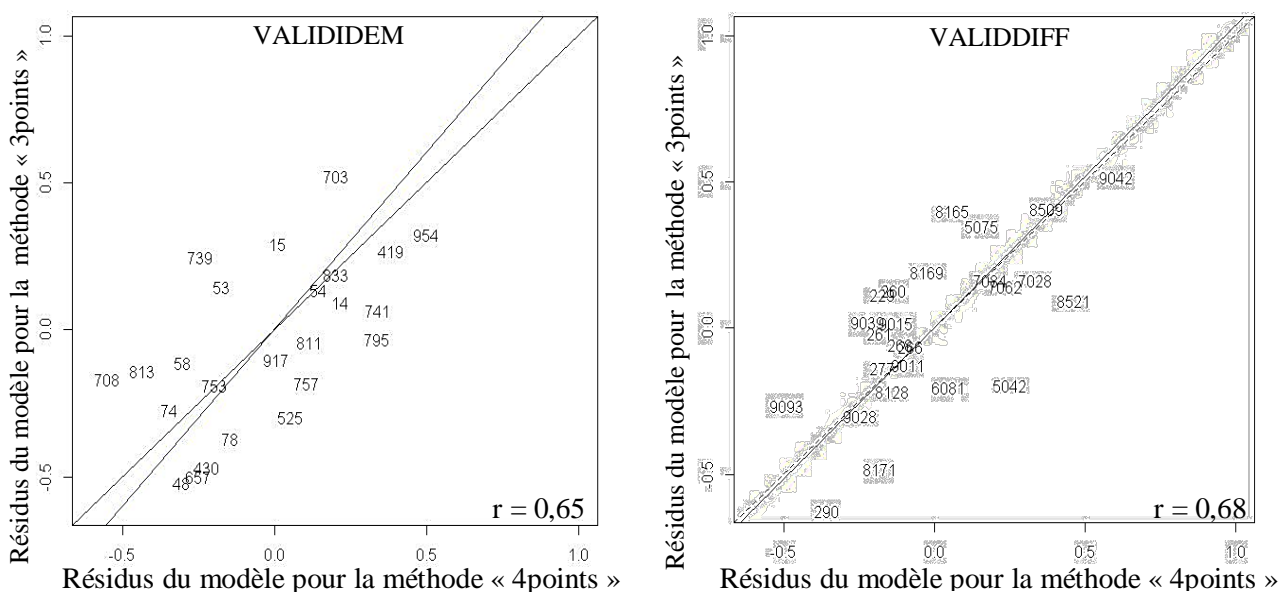


Figure 20 : Comparaison des résidus de prédiction de la nec3d des 2 méthodes de pointage pour les 2 populations de validation

r : coefficient de corrélation entre les 2 variables représentées sur le graphique.

n'est pas le même que celui optimisé sur la qualité de la calibration.

Le modèle conservé pour l'estimation de la nec3d est celui optimisé sur la qualité de la calibration contenant 55 composantes. En validation externe, le modèle prédit la NEC observée avec une erreur quadratique moyenne comprise entre 0,31 et 0,37 pour les 2 populations de validations et les 2 méthodes de pointage. L'erreur quadratique moyenne réalisée sur la prédiction par le modèle de la NEC de la population « valididem » est inexpliquée à 95% pour les 2 méthodes de pointage, les 5% restants sont répartis dans le biais moyen et le biais de pente (tableau 6 et annexe IV). La répartition de l'erreur quadratique moyenne (MSE) pour la population « validdiff » présente un biais de pente plus important : près de 70% de l'erreur est inexpliquée et 29% est expliquée par la pente, indépendamment de la méthode de pointage.

Indépendamment de la nature de la population de validation, l'erreur quadratique moyenne de l'estimation de la NEC3D associée au pointage « 3points » est un peu plus faible que l'erreur associée au pointage « 4points ». De plus, la MSE est similaire pour les 2 populations de validation, l'appartenance ou non des vaches à la population de calibration n'a donc pas d'influence sur la qualité prédictive du modèle (tableau 6). L'existence d'une relation entre les résidus des modèles pour chacune des populations de validation permet de voir que les erreurs de prédiction sont largement liées à d'autres facteurs que la méthode de pointage. Les résidus des modèles des 2 méthodes de pointage sont significativement corrélés avec des coefficients de 0,65 et 0,68 respectivement pour les estimations faites sur la population « valididem » et « validdiff » (figure 20) et l'axe principal de ces relations est proche de la première bissectrice.

En conclusion, les résultats de validation ne permettent pas de qualifier une des 2 méthodes de pointage comme étant de meilleure qualité prédictive que l'autre. Néanmoins, étant donné que les erreurs de prédiction réalisées sur les 2 populations de validation sont les plus faibles pour la méthode « 3points », c'est cette méthode que nous avons choisi de conserver pour la suite des analyses, notamment en caractérisant l'erreur liée au pointage, indépendamment de l'erreur liée à la mesure de la NEC.

3. Une méthode plus répétable et reproductible que la NEC

La répétabilité et la reproductibilité permettent de qualifier l'erreur réalisée à chaque mesure, indépendamment de l'erreur de mesure de la NEC, liée d'une part au pointage en ce qui concerne la répétabilité et d'autre part au système d'acquisition en ce qui concerne la reproductibilité. L'évolution de l'écart-type de répétabilité, σ_r , calculé pour les différents modèles au cours de la sélection « pas à pas » permet d'évaluer la pertinence du nombre de composantes dans le modèle (5) conservée à l'issue de la calibration.

L'écart-type de répétabilité atteint son minimum, $\sigma_r = 0,068$, à la 28^{ème} étape de sélection « pas à pas », c'est-à-dire pour un modèle contenant 28 composantes. Le modèle (5) sélectionné à l'issue du stepwise n'est pas le modèle offrant la plus grande répétabilité de mesure, $\sigma_r = 0,075$. L'erreur liée au pointage à chaque mesure de nec3d s'élève à 0,075, c'est-à-dire un coefficient de variation de 2,91%. La reproductibilité de la nec3d mesurée sur population « reproductibilité » vaut $\sigma_R = 0,10$, ce qui se traduit en un coefficient de variation de 3,95%. A chaque estimation de la NEC par la méthode « nec3d », la mesure contient une

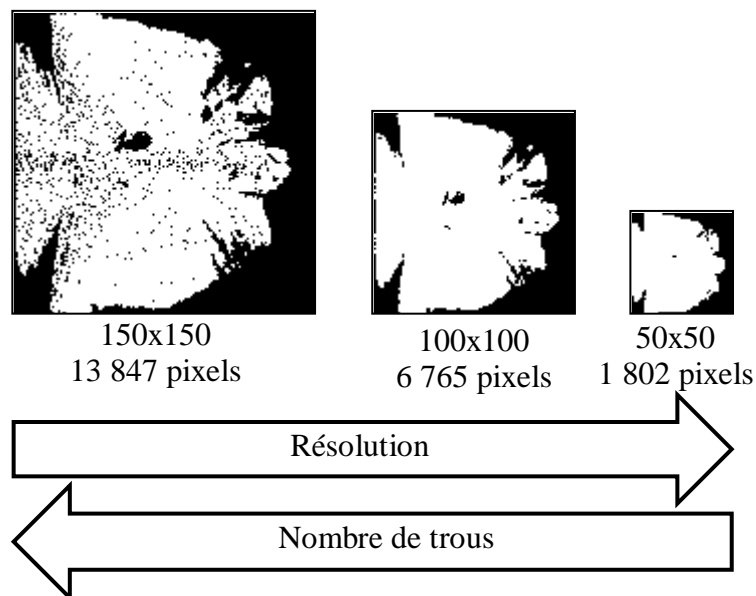


Figure 21 : Comparaison des résolutions et du nombre de pixels renseignés pour les masques obtenus pour 3 résolutions: 150x150, 100x100, 50x50

Tableau 7: Résultats de validation des modèles associés aux 3 résolutions: 150x150, 100x100 et 50x50 pour les 2 populations de validation, valididem et validdiff

| Population | Résolution | MSE | Bm (%) | Bp (%) | Bi (%) | R ² | RMSE |
|------------|------------|-------|--------|--------|--------|----------------|------|
| valididem | 150x150 | 0,094 | 0,07 | 4,49 | 95,44 | 0,91 | 0,31 |
| | 100x100 | 0,096 | 1,67 | 2,24 | 96,09 | 0,91 | 0,31 |
| | 50x50 | 0,094 | 0,19 | 0,27 | 99,53 | 0,91 | 0,31 |
| validdiff | 150x150 | 0,100 | 0,82 | 28,83 | 70,36 | 0,80 | 0,32 |
| | 100x100 | 0,094 | 2,77 | 26,87 | 70,36 | 0,81 | 0,31 |
| | 50x50 | 0,103 | 2,66 | 26,42 | 70,93 | 0,79 | 0,32 |

MSE : mean squared error = erreur quadratique moyenne / $RMSE = \sqrt{MSE}$

Bm : biais de la moyenne (% de la MSE) / *Bp* : biais de la pente (% de la MSE)

Bi : biais inexpliqué (% de la MSE)

erreur de 0,10 qui est indépendante de l'erreur de mesure par les notateurs de la NEC. L'écart entre la reproductibilité, qui comptabilise à la fois l'erreur liée au pointage et l'erreur de mesure du système d'acquisition, et la répétabilité est faible. Il est donc possible qu'une partie importante de l'erreur commise par la méthode « nec3d » vient du pointage des surfaces 3D.

4. La résolution du masque a peu d'impact sur la valeur prédictive de la méthode « nec3D ».

La grille d'échantillonnage de l'image pour aboutir au masque final définit le nombre de pixels conservés par surface 3D pour l'ACP : plus la résolution de la grille diminue, plus le nombre de pixels par masque est faible. Le passage d'une grille de dimensions 150x150 à une grille de dimensions 100x100 réduit le nombre de pixels réel du masque d'un facteur 2,05 et d'un facteur 7,7 pour une grille de dimensions 50x50. Les masques obtenus pour chacune des 3 résolutions témoignent d'une fréquence de reconstruction par le traitement d'image plus élevée des zones d'ombre quand la résolution diminue (figure 21). Les surfaces 3D sont donc plus complètes avec une résolution plus basse, à 50x50 mais contiennent moins d'information que pour des résolutions plus grandes, à 150x150.

La comparaison des 3 grilles permet de déterminer l'influence de la résolution du masque, donc de la quantité d'information valorisée par surface 3D, sur la qualité de l'estimation de la NEC par la méthode « nec3d » en termes de validation externe et de fidélité de la méthode. Seule la méthode de pointage « 3points » a été utilisée pour cette comparaison.

Les ACP réalisées sur la population de calibration pour les 3 résolutions choisies définissent des composantes principales dont l'inertie cumulée augmente plus rapidement avec des résolutions élevées. Pour expliquer 95% de la variabilité des surfaces 3D dans la population de calibration, le nombre de composantes principales nécessaires augmente de 13 à 19 respectivement pour les résolutions de masque de 150x150 et 50x50. En parallèle, l'évolution de la part de la variabilité de la NEC expliquée par les modèles de chaque étape de la sélection « pas à pas » relève la même tendance : pour atteindre un coefficient de détermination (R^2) de 95%, il faut pour une résolution croissante, 16, 18 et 15 composantes dans le modèle.

Le meilleur modèle sélectionné par la méthode « pas à pas » contient dans les 3 cas 55 composantes dont l'ordre d'apparition dans le modèle est différent de celui des composantes principales dans l'ACP et diffère pour chacune des résolutions. Néanmoins, les 2 premières composantes de l'ACP figurent dans le même ordre et avec le même rang dans chacun des modèles obtenus. La définition de ces 2 composantes semble être indépendante de la résolution du masque. Les modèles associés à chaque résolution de masque sont caractérisés par des qualités de calibration similaires avec un R^2 ajusté proche de 1 et une erreur résiduelle comprise entre 1.10^{-4} et 1.10^{-3} .

La qualité des 3 modèles diffère faiblement en termes de qualités prédictives sur les populations de validation (tableau 7) externe. L'étape de validation est en faveur du modèle associé à la résolution la plus haute pour la population « valididem » alors qu'elle est en faveur du modèle associé à la résolution intermédiaire pour la population « validdiff ». Quelle que soit la population de validation utilisée, l'augmentation de la résolution du masque s'accompagne d'une diminution de la part inexpliquée de l'erreur quadratique moyenne de

Tableau 8: Corrélation entre les résidus issus des prédictions réalisées par les modèles associés aux résolutions (150x150, 100x100 et 50x50) pour les 2 populations de validation (valididem et validdiff)

| | 150x150 | 100x100 | 50x50 |
|---------|---------|--------------------|--------------------|
| 150x150 | 1 | 0,96 / <u>0,97</u> | 0,89 / <u>0,96</u> |
| 100x100 | / | 1 | 0,90 / <u>0,99</u> |
| 50x50 | / | / | 1 |

Texte : coefficient de corrélation pour la population « valididem »

Texte : coefficient de corrélation pour la population « validdiff »

Tableau 9: Comparaison de la répétabilité et de la reproductibilité de la nec3d pour les résolutions de masque: 150x150, 100x100 et 50x50

| Résolution | σ_r | σ_R | CVr | CVR |
|------------|------------|------------|------|------|
| 150x150 | 0,07 | 0,10 | 2,91 | 3,95 |
| 100x100 | 0,07 | 0,11 | 2,59 | 4,30 |
| 50x50 | 0,09 | 0,13 | 3,65 | 4,90 |

σ_r : écart-type de répétabilité / σ_R : écart-type de reproductibilité

CVr : coefficient de variations de répétabilité / CVR : coefficient de variations de reproductibilité

99,5% à 95,4% au profit de l'erreur expliquée par un biais de pente. De plus, pour la population « valididem », les résidus associés à la résolution 50x50 sont corrélés à 89% avec les résidus des 2 autres résolutions, alors qu'ils sont corrélés à 96% pour les 2 autres résolutions et à plus de 96% pour les prédictions réalisées sur la population « validdiff » (tableau 8 et annexe V).

La résolution des masques est cependant plus importante en termes de fidélité des mesures. En effet, la diminution de la résolution crée une perte de répétabilité et de reproductibilité des modèles (tableau 9). Néanmoins, c'est principalement la résolution 50x50 qui semble dégrader la qualité, la résolution 100x100, bien qu'offrant une reproductibilité intermédiaire à la résolution 150x150 et 50x50, est caractérisée par la meilleure répétabilité : $\sigma_r = 0,066$ et une faible diminution de la reproductibilité.

En conclusion, une perte de résolution au niveau du masque n'a pas d'influence importante en termes de qualité de calibration et de validation des modèles d'estimation de la NEC par la nec3d mais engendre une perte de répétabilité et de reproductibilité.

DISCUSSION

Le projet visait à définir une méthode innovante pour estimer la NEC des vaches laitières à partir de surfaces 3D de la partie dorsale postérieure de la vache. Une méthode d'acquisition semblable a été utilisée avec un capteur Kinect (Microsoft®) par Kawasue et al. (2012) pour mesurer automatiquement les mesures de gabarit des vaches. Le système d'acquisition, bien que gérant une dimension supplémentaire par rapport au système analysant les images 2D, est aussi simple et rapide que le système d'acquisition d'images 2D (Bewley et al. 2008, Halachmi et al. 2008, Negretti et al. 2008). De plus, le capteur et le système d'acquisition choisis ne sont ni sensibles aux variations de luminosité comme pour Bewley et al. (2008), ni victimes de la saleté projetée par le nettoyage de la station de pesée comme pour Azzaro et al. (2011). Cette différence de sensibilité se traduit en termes de qualité d'images : le système d'acquisition adopté dans cette étude n'a dans le pire des cas que réalisé 70% de surfaces de bonne qualité alors que pour Bewley et al. (2008) ce chiffre s'élève en moyenne à 23% et à 2,75% pour Azzaro et al. (2011). L'approche de transformation des acquisitions par ACP qui a été retenue semble performante pour prédire l'état corporel, sans pour autant appliquer une fonction supplémentaire aux données - l'ACP est alors dite à noyau – ce qui est souvent utilisé pour analyser des formes (Cootes et al., 1994 ; Sahbi 2007, cité par Azzaro et al. 2011). La méthodologie choisie pour estimer la NEC par le nec3d passe par 2 phases : une phase de traitement de l'information 3D brute et de calibration de l'information 3D sur la NEC et une phase de qualification de la méthode utilisée. C'est le choix des paramètres caractérisant chacune des étapes du traitement des surfaces 3D et l'étape de calibration qui conditionnent la qualité des estimations de NEC. Nous nous sommes plus particulièrement concentrés sur l'influence de la localisation des 4 points du pointage et de la résolution du masque.

1. L'influence du choix du paramétrage sur la nec3d

Le pointage est l'étape essentielle dans toute méthode équivalente d'imagerie. Ici, l'identification a été réalisée manuellement avec une vérification manuelle de la pertinence des points enregistrés pour chaque forme. C'est l'étape la plus longue et la plus pénible de la méthode : en moyenne l'identification dure 3 minutes et la vérification dure 20 secondes. Les travaux d'Azzaro et al. (2011) et Bewley et al. (2008) étaient basés sur une identification à 23 points anatomiques. En réduisant le nombre de points à 4, nous réalisons donc un gain de temps considérable, mais en perdons dans le processus de vérification des 4 points enregistrés par surface, étant donné que pour les travaux antérieurs, cette étape était automatisée. L'identification des 4 mêmes points sur toutes les surfaces 3D, permet de passer du repère du « capteur » à celui de l'« animal ». Les 2 pointages, « 4points » dont les 2 points postérieurs identifient les pointes de fesses et « 3points » dont les 2 points postérieurs identifient la base du sacrum, ont été réalisés sur les mêmes acquisitions : celles de la population de calibration. Par conséquent, la transformation appliquée aux surfaces 3D devrait aboutir à des masques similaires. Les résultats mettent en évidence une compression dans le sens de la longueur des surfaces pointées avec la méthode « 4points » par rapport à la méthode « 3points ». Cette compression pourrait provenir d'une différence de rapport d'homothétie entre les deux méthodes qui pourrait être corrigé si nécessaire. C'est probablement cette compression qui explique la réduction du nombre de trous dans le masque associé au pointage « 4points » et donc un nombre de pixels par masque plus important. Les résultats de calibration et de validation ne montrent aucune différence significative entre les 2 méthodes de pointage, par conséquent, le surplus d'informations apporté par le pointage « 4points » ne véhicule pas d'information capitale associée aux variations de NEC. Un pointage basé sur 3 points identifiant les pointes de fesses et la base du sacrum suffirait pour fournir des estimations de qualité de la NEC à partir de la nec3d et permettrait de supprimer l'étape de calcul pour établir le plan de référence.

En diminuant la résolution, les masques n'ont quasiment plus de zones d'ombres et sont par conséquent plus complets. Cette différence ne ressort ni au niveau de la qualité de calibration, ni au niveau de la qualité de validation, étant donné que les résultats sont similaires pour les 3 résolutions. Néanmoins, bien que le masque 50x50 soit plus complet que le 100x100, lui-même plus complet que le masque 150x150, il contient moins d'informations et surtout aboutit à des mesures de NEC moins répétables et moins reproductibles que des estimations faites sur des résolutions plus hautes, 100x100, lui-même moins reproductible que le masque 150x150. Le masque de résolution 150x150 est donc le masque le plus adapté pour des mesures dédiées à l'analyse de la cinétique de la nec3d des vaches, mais le masque 50x50 suffit pour estimer la nec3d ponctuellement avec une qualité équivalente aux résolutions élevées. L'utilisation de faible résolution d'acquisition offre la possibilité de recourir à des technologies moins coûteuses et moins contraignantes en termes de qualité d'acquisition que la technique actuelle. La technique du « **one-shot**¹² » fournit des acquisitions de résolution plus faible, mais est moins chère que le prototype actuel et est moins sensible au mouvement

¹² La technique du « one-shot » est basée sur une unique prise de vue de l'animal et non sur la superposition de 90 prises de vue.

Tableau 10: Comparaison de la répétabilité et de la reproductibilité des mesures de NEC par les notateurs et des estimations de NEC par la nec3d

| méthode | σ_r | σ_R | CV_r | CV_R |
|---------|------------|------------|--------|--------|
| nec3d | 0,07 | 0,10 | 2,91 | 3,95 |
| NEC | 0,21 | 0,28 | 10,05 | 13,44 |

σ_r : écart-type de répétabilité / σ_R : écart-type de reproductibilité

CV_r : coefficient de variations de répétabilité en %

CV_R : coefficient de variations de reproductibilité en %

des vaches, ne nécessitant pas l'immobilisation des vaches pour avoir des acquisitions de qualité.

A l'issue de cette transformation des surfaces 3D brutes, chacune des surfaces est réduite aux informations de la surface projetées sur le masque obtenu. Cette information, projetée dans le repère de l'ACP, ne contient plus que 55 valeurs par vache qui sont utilisées pour calibrer ces données sur la NEC par régression linéaire multiple. Le meilleur modèle d'estimation obtenu avec la sélection « pas à pas » conserve 55 composantes principales, c'est-à-dire le maximum de composantes que le modèle peut contenir. Néanmoins l'évolution de l'écart-type de répétabilité associé aux modèles de chaque étape de la sélection « pas à pas » met en évidence un minimum bien avant l'incorporation des 55 composantes dans le modèle. Le modèle à 55 composantes principales est sans doute surparamétré, les 27 dernières composantes ajoutées dans le modèle augmentant l'erreur de répétabilité. Ce résultat est similaire pour l'évolution de l'erreur quadratique moyenne des prédictions réalisées par les modèles issus des étapes de la sélection « pas à pas » en validation externe. Ces 27 composantes servent probablement dans le modèle à expliquer les spécificités de certaines images ou de certains individus, mais ne portent pas de valeur générique. Le modèle à 28 composantes est moins bien ajusté à la population de calibration ($R^2= 0,98$) que le modèle complet, mais fournit d'aussi bon résultats de validation externe et améliore la répétabilité des mesures. Pour vérifier cette hypothèse, une analyse d'influence de chacune des composantes permettrait de vérifier qu'aucune composante n'est influencée que par quelques individus et une **ridge regression**¹³ permettrait d'analyser la robustesse des coefficients du modèle. Ces étapes complémentaires seraient utiles lors de l'établissement de l'équation retenue pour la calibration définitive de la méthode.

Les résultats, tels qu'ils ont été présentés, peuvent aider à améliorer la méthode en choisissant les paramètres optimaux pour répondre aux objectifs fixés. Ils permettent cependant déjà de donner des éléments de réponse par rapport aux attentes fixées pour le projet en termes de caractérisation de la méthode. Néanmoins, il faut garder à l'esprit que les résultats ont été obtenus en sélectionnant des acquisitions de bonne qualité. Les critères de cette sélection pour permettre de l'automatiser restent à définir.

2. Une méthode répétable et reproductible

La méthode d'estimation de la NEC basée sur un pointage « 3points » avec un masque de résolution 150x150 est très répétable et reproductible : chaque mesure de nec3d possède un coefficient de variation de 3,95% relatif à la méthode, indépendamment de l'erreur incluse dans les mesures de NEC. En effet, en calibrant la nec3d sur la NEC, chaque mesure comprend l'erreur caractéristique des mesures de NEC qui reste une méthode subjective et peu précise. A l'installation expérimentale de Méjusseume, les mesures de NEC réalisées avec 3 notateurs dans le cadre de cette étude présentent un écart-type de répétabilité de 0,21 et de reproductibilité de 0,28 (tableau 10). La nec3d est donc 3 fois plus fidèle que la NEC. La

¹³ Ridge régression ou régression pseudo-orthogonalisée : cette régression impose une longueur maximale à la norme du vecteur des coefficients du modèle de régression. En faisant varier cette longueur maximale, la régression permet d'identifier les variables instables, dont les coefficients sont instables et tendent vers zéro et dont les coefficients sont faibles (Tomassone et al., 1983).

répétabilité de 0,07 de la nec3d est associée à l'erreur liée au pointage manuel des acquisitions. En rendant le pointage automatique, l'identification des points ne serait plus sujette à la subjectivité humaine. L'écart-type de répétabilité de la méthode s'en verrait probablement annulé et par conséquent, l'écart-type de reproductibilité pourrait s'en trouver diminué si l'automatisation du pointage est performante.

3. L'imagerie 3D : une technique avec des perspectives d'avenir

La technique d'estimation de la NEC présentée dans cette étude présente de nettes améliorations en termes de qualité d'estimation. Azzaro et al. (2011) ont comparé la qualité des validations de leur méthode basée sur une ACP à noyau, de la méthode de Bewley et al. (2008) et de la méthode d'Halachmi et al. (2011). Par comparaison, l'erreur de prédiction du modèle de régression (5) est équivalente à celle de la technique basée sur l'ACP à noyau qui parmi les 3 méthodes comparées par Azzaro et al. (2011) est celle qui possède l'erreur la plus faible. En plus d'avoir une erreur de prédiction équivalente à celle du meilleur modèle obtenu dans les travaux précédents, la nec3d renseigne sur la fidélité de la méthode, avec une répétabilité et reproductibilité élevées qui n'ont pas été calculées dans les travaux récents (Bewley et al., 2008 ; Halachmi et al., 2008 ; Azzaro et al., 2011).

L'imagerie 3D appliquée à l'estimation de la NEC des vaches laitières fournit des résultats prometteurs en termes de nouvelles méthodes de mesure de l'état des réserves corporelles. En effet, les erreurs de prédiction et les écart-types de répétabilité et de reproductibilité sont nettement plus faibles que ceux de la NEC. De plus, l'écart-type de reproductibilité de la NEC est du même ordre de grandeur que l'erreur quadratique moyenne des prédictions réalisées par le modèle (5) sur une population de validation externe (RMSE = 0,31 pour la population « valididem » et 0,32 pour « validdiff »). Néanmoins, ces prédictions ont été faites sur une technique de cross-validation par LOOCV alors que les prédictions réalisées avec le modèle (5) ont été faites sur des populations de validation complètement différentes de la population de calibration. L'erreur introduite par la NEC semble incluse dans la RMSE de la nec3d sur les populations de validation externe. Si le choix de la NEC comme référence peut se justifier parce que c'est un indicateur largement répandu sur le terrain, la NEC ne présente pas les qualités d'un bon « **gold standard**¹⁴ ». Une calibration de la nec3d sur une méthode de mesure de l'état corporel plus précise et plus objective que la NEC, telle que l'épaisseur de gras sous-cutané par échographie (Schröder et Staufenbiel, 2006), diminuerait probablement l'erreur quadratique moyenne des prédictions réalisées sur une population de validation externe. Ainsi, en complétant l'automatisation du pointage par une calibration sur une méthode de référence plus précise que la NEC, la nec3d gagnerait en qualité prédictive sur des acquisitions externes à la population de calibration.

L'information 3D a été valorisée pour estimer la NEC des vaches laitières, mais pourrait servir pour évaluer d'autres phénotypes. En pratique, la méthodologie est identique à celle employée pour la nec3d jusqu'à la calibration où le modèle linéaire serait ajusté sur ce nouveau critère phénotypique à la place de la NEC. Une telle calibration nécessiterait probablement de revoir la définition de la zone de surface 3D conservée pour l'analyse. Sous

¹⁴ Un gold standard est un test qui fait référence dans un domaine pour établir la validité d'un fait.

réserve d'automatisation du traitement de l'image, cette méthode de notation par imagerie 3D peut présenter un compromis optimal en étant aussi précise que des méthodes longues et coûteuse d'échographie (Schröder et Staufenbiel, 2006) ou de mesure des adipocytes (Robelin et al., 1989). Mais d'autres phénotypes tels que le poids vif ou les autres mesures liées à la morphologie des animaux pourraient être approchés par ce type d'imagerie 3D en élargissant les angles de vue (Tasdemir et al., 2011 ; Kawasue et al., 2012). Pour le poids vif par exemple, étant donné qu'il est associé au gabarit et à l'état corporel de la vache, pourrait être estimé par imagerie 3D en ajoutant une information relative au gabarit de la vache mesurable par la dimension du trapèze des 4 points du repère par exemple.

CONCLUSION

Le projet a permis de construire une méthode basée sur des techniques d'imagerie 3D pour estimer objectivement et avec une meilleure reproductibilité la NEC des vaches laitières à partir de la forme des surfaces 3D de la partie postérieure dorsale des vaches. En plus de présenter et tester un nouvel indicateur, la nec3d, la démarche du projet a permis de valider et de qualifier la méthode.

La méthodologie développée a permis de calibrer parfaitement la NEC du jeu d'essai à partir du modèle de traitement de l'image 3D. Les premiers résultats obtenus en termes de qualité de la méthode sont bons avec une reproductibilité de la mesure présentant un coefficient de variation de l'ordre 4%. L'erreur de prédiction avec la validation externe est de l'ordre 0,32 point de note d'état, ce qui est voisin de l'erreur de notation par la NEC. De plus, étant donné que cette erreur est la même pour une population contenant des acquisitions de vaches incluses ou non dans la population de calibration, la population de calibration choisie contient assez de variabilité pour que les prédictions en validation externe soient de bonne qualité. Ces résultats sont donc très encourageants et motivent l'espoir de pouvoir rapidement remplacer la notation réalisée par l'homme par cette technique plus automatique.

Néanmoins étant donné que la méthode de référence utilisée est la NEC et que l'étape d'identification des 4 points de référence est manuelle, la nec3d demeure encore un peu subjective. L'automatisation du pointage se présente comme une étape indispensable pour rendre la méthode plus objective. De plus, par ce biais, la méthode est totalement répétable et permettra également de gagner en reproductibilité. La principale difficulté sera de caractériser précisément la localisation des points de référence. Cette nouvelle technique est d'autant plus prometteuse qu'une calibration sur un indicateur plus objectif et plus précis que la NEC, tel que l'épaisseur des tissus sous-cutanés mesuré par échographie, améliorerait la qualité de la méthode. Des étapes complémentaires sont donc à réaliser pour faire définitivement de cette méthode, une méthode de référence.

La qualité des résultats obtenus avec l'imagerie 3D encourage l'utilisation de cette nouvelle technologie en élevage. Les coûts relativement faibles des technologies mobilisées et les gains de temps possibles permis par l'automatisation de cette notation ouvrent la possibilité de disposer enfin d'une méthode de phénotypage à haut débit de l'état corporel. Cette innovation ouvre notamment à la recherche et à l'élevage de nouvelles perspectives à la fois pour l'étude de nouveaux caractères et leur sélection, mais aussi pour l'élevage de précision.

RÉFÉRENCES BIBLIOGRAPHIQUES

1. Sites internet

[1] : valorisation de la NEC comme outil de gestion du troupeau : *Dairy NZ Body Condition scoring : The reference guide for New Zealand dairy farmers*, 2012. Consulté le 30/08/2013 sur : <http://www.dairynz.co.nz/page/pageid/2145864561?resourceId=703>

[2] : définition et calcul de la répétabilité et de la reproductibilité intra-laboratoire : P. Jost, *Statistiques à l'usage des ingénieurs et des techniciens*. Consulté le 29/07/2013 sur : http://infochimie.u-strasbg.fr/master/Cours_stat_pdf/ANOVA.PDF

[3] : régression linéaire ajustée par la technique des moindres rectangles : Borcard D., *Régression linéaire simple de modèle II*. Consulté le 30/08/2013 sur : http://biol09.biol.umontreal.ca/bio2042/Axe_majeur.pdf

2. Articles scientifiques

Azzaro, G., Caccamo, M., Ferguson, J.D., Battiato, S., Farinella, G.M., Guarnera, G.C., Puglisi, G., Petriglieri, R. and Licitra, G. (2011). Objective estimation of body condition score by modeling cow body shape from digital images. *Journal of Dairy Science*, 94, 2126-2137.

Bazin, S. and Augeard, P. (1984). *Grille de notation de l'état d'engraissement des vaches pie noires*. Institut technique de l'Élevage bovin pour le Réseau national d'Experimentation et des Démonstrations en Elevage bovin, 31 pp.

Berry, D. P., Buckley, F., Dillon, P., Evans, R. D., Rath, M. and Veerkamp, R. F. (2003). Genetic Relationships among Body Condition Score, Body Weight, Milk Yield, and Fertility in Dairy Cows. *Journal of dairy science*, 86, 2193-2204.

Bewley, J. M., Peacock, A.M., Lewis, O., Boyce, R.E., Roberts, D.J., Coffey, M.P., Kenyon, S.J. and Schutz, M.M. (2008). Potential for Estimation of Body Condition Scores in Dairy Cattle from Digital Images. *Journal of Dairy Science*, 91, 3439-3453.

Campanile, G., De Filippo, C., Di Palo, R., Taccone, W. and Zicarelli, L. (1998). Influence of dietary protein on urea levels in blood and milk of buffalo cows. *Livestock Production Science*, 55, 135-143.

Coffey, M. P., Simm, G., Hill, W. G. and Brotherstone, S. (2003). Selection for energy balance in dairy cows. *Interbull Bulletin*, 107-111.

Cootes, T. F., Taylor, C. J., Cooper, D. H. and Graham, J. (1992). Training Models of Shape from Sets of Examples. In Hogg, D. and Boyle, R. (Eds.) *British Machine Vision Conference 1992*, Springer, London, p. 9-18..

Dechow, C. D., Rogers, G. W. and Clay, J. S. (2002). Heritability and Correlations Among Body Condition Score Loss, Body Condition Score, Production and Reproductive Performance. *Journal of dairy science*, 85, 3062-3070.

Earle, D. F. (1976). A guide to scoring dairy cow condition. *Journal of Agriculture (Victoria)*, 74, 228-231.

- Edmonson, A. J., Lean, I. J., Weaver, L. D., Farver, T. and Webster, G.** (1989). A body condition scoring chart for Holstein dairy cows. *Journal of dairy science*, 72, 68-78.
- Feinberg, M.** (2010). Principes et vocabulaire pour la validation des méthodes. In Feinberg, M. et Laurentie, M. (Eds) *Validation des méthodes d'analyse quantitative par le profil d'exactitude, numéro spécial des cahiers techniques de l'INRA*, INRA Délégation au Partenariat avec les Entreprises, Paris, p. 13-25.
- Ferguson, J. D., Azzaro, G. and Licitra, G.** (2006). Body Condition Assessment Using Digital Images. *Journal of Dairy Science*, 89, 3833-3841.
- Ferguson, J. D., Galligan, D. T. and Thomsen, N.** (1994). Principal Descriptors of Body Condition Score in Holstein Cows. *Journal of dairy science*, 77, 2695-2703.
- Gillund, P., Reksen, O., Gr+Åhn, Y. T. and Karlberg, K.** (2001). Body Condition Related to Ketosis and Reproductive Performance in Norwegian Dairy Cows. *Journal of dairy science*, 84, 1390-1396.
- Halachmi, I., Polak, P., Roberts, D. J. and Klopčič, M.** (2008). Cow Body Shape and Automation of Condition Scoring. *Journal of Dairy Science*, 91, 4444-4451.
- Hastie, T., Tibshirani, R. and Friedman, J.** (2009). *The Elements of Statistical Learning: Data Mining, Inference, and Prediction (Second Edition)*. Springer Verlag, New-York, p. 547-550.
- Kawasue, K. I. K. U., Ikeda, T. A. K. A., Tokunaga, T. A. D. A. and Harada, H. I. R. O.** (2012). Three-Dimensional Measurement System for Japanese Black Cattle Using Computer Vision. In Pavelkova, D., Strouhal, J. and Pasekova, M. (Eds.) *Advances in environment, biotechnology and biomedicine*, WSEAS Press, p. 229-234.
- Krukowski, M.** (2009). Automatic Determination of Body Condition Score of Dairy Cows from 3D Images. In : Royal Institute of Technology, School of Computer Science and Communication.
- Lafaye de Micheaux, P. L., Drouilhet, R. and Liquet, B.** (2011). *Le logiciel R: Maîtriser le langage - Effectuer des analyses statistiques*. Springer Verlag, France, p. 404-407.
- Lowman, B. G., Scott, N. A. and Somerville, S. H.** (1976). Condition scoring of cattle. *East of Scotland College of Agriculture, Bulletin No. 6, Edinburgh, UK*.
- Luginbühl, T.** (2012). Régularisation de surface 3D par une approche fondée sur un modèle. Thèse de doctorat de l'Université Pierre et Marie Curie, Paris, Ecole doctorale informatique, télécommunications et électronique (EDITE), 196 pp.
- Macdonald, K. and Roche, J.** (2004). Condition Scoring Made Easy. Condition scoring dairy herds. Dexcel Publishing, Hamilton, New Zealand. 36 pp.
- Negretti, P., Bianconi, G., Bartocci, S., Terramoccia, S. and Verna, M.** (2008). Determination of live weight and body condition score in lactating Mediterranean buffalo by Visual Image Analysis. *Livestock Science*, 113, 1-7.

- Pryce, J. E. and Harris, B. L.** (2006). Genetics of Body Condition Score in New Zealand Dairy Cows. *Journal of dairy science*, 89, 4424-4432.
- Pryce, J. E., Coffey, M. P. and Brotherstone, S.** (2000). The Genetic Relationship between Calving Interval, Body Condition Score and Linear Type and Management Traits in Registered Holsteins. *Journal of dairy science*, 83, 2664-2671.
- Robelin, J., Barboiron, C. and Jailler, R.** (1982). Relation entre l'espace de diffusion de l'eau lourde mesurée in vivo et le volume hydrique corporel des bovins en croissance. *Reproduction Nutrition Development*, 22, 65-73.
- Robelin, J., Chilliard, Y. and Agabriel, J.** (1989). Estimation of body lipids and proteins of Holstein Charolaise and Limousine cows by dilution technique and adipose cell size. *Symposium on Energy Metabolism of Farm Animals, Lunteren (Netherlands), 18-24 Sep 1988*.
- Roche, J. R. and Berry, D. P.** (2006). Periparturient Climatic, Animal, and Management Factors Influencing the Incidence of Milk Fever in Grazing Systems. *Journal of dairy science*, 89, 2775-2783.
- Roche, J. R., Friggens, N. C., Kay, J. K., Fisher, M. W., Stafford, K. J. and Berry, D. P.** (2009). Invited review: Body condition score and its association with dairy cow productivity, health, and welfare. *Journal of Dairy Science*, 92, 5769-5801.
- Ruegg, P. L. and Milton, R. L.** (1995). Body Condition Scores of Holstein Cows on Prince Edward Island, Canada: Relationships with Yield, Reproductive Performance, and Disease. *Journal of dairy science*, 78, 552-564.
- Sahbi, H.** (2007). Kernel PCA for similarity invariant shape recognition. *Neurocomputing*, 70, 3034-3045.
- Schröder, U. J. and Staufenbiel, R.** (2006). Invited Review: Methods to Determine Body Fat Reserves in the Dairy Cow with Special Regard to Ultrasonographic Measurement of Backfat Thickness. *Journal of dairy science*, 89, 1-14.
- Szabo, C., Babinszky, L., Versteegen, M. W. A., Vangen, O., Jansman, A. J. M. and Kanis, E.** (1999). The application of digital imaging techniques in the in vivo estimation of the body composition of pigs: a review. *Livestock Production Science*, 60, 1-11.
- Tasdemir, S., Urkmez, A. and Inal, S.** (2011). Determination of body measurements on the Holstein cows using digital image analysis and estimation of live weight with regression analysis. *Computers and Electronics in Agriculture*, 76, 189-197.
- Tomassone, R., Lesquoy, E. and Millier, C.** (1983). *La régression: nouveaux regards sur une ancienne méthode statistique*. INRA Elsevier, Paris, p. 153-158..
- Wallach, D.** (2006). Evaluating crop models. In Wallach, D., Makowski, D. and Jones, J. W. (Eds.) *Working with dynamic crop models*. Elsevier Science, p. 17-25.

Waltner, S. S., McNamara, J. P., Hillers, J. K. and Brown, D. L. (1994). Validation of indirect measures of body fat in lactating cows. *Journal of dairy science*, 77, 2570-2579.

Wildman, E. E., Jones, G. M., Wagner, P. E., Boman, R. L., Troutt Jr, H. F. and Lesch, T. N. (1982). A dairy cow body condition scoring system and its relationship to selected production characteristics. *Journal of dairy science*, 65, 495-501.

ANNEXES

Annexe I : Grille de notation de l'état d'engraissement des vaches Pie Noires, d'après le Bazin (1984)

| DESCRIPTION DES CRITERES | |
|--|---|
| A – QUELS CRITERES OBSERVER POUR ATTRIBUER LA NOTE "ARRIERE" ? | |
| <p>Nous allons décrire les trois critères visuels qui permettent d'attribuer la note ARRIERE. Ils portent sur l'observation de :</p> <ul style="list-style-type: none"> – la base de la queue et la pointe des fesses, – le ligament sacrotubéral et le détroit caudal, – la ligne du dos. <p>Nous compléterons, dans certains cas, par des informations sur le maniement des abords.</p> | |
| 1 - Base de la queue, pointe de la fesse. | |
| <p>On gardera toujours en mémoire que suivant le rameau (Holstein ou Frison) et suivant les souches, la conformation osseuse et musculaire de cette région est très variable.</p> | |
| NOTE 5 | La queue est "noyée" dans un rond de tissu gras centré à la partie supérieure de la vulve, descendant largement sous la pointe de la fesse. En rameau Holstein les masses autour de la queue, bien qu'abondantes, peuvent la laisser dégagée. |
| NOTE 4 | Le "rond" n'existe plus, il reste, autour des boules marquées à la pointe des fesses, un croissant aplati à concavité vers le haut correspondant à un "débordement" des boules aux pointes des fesses. |
| NOTE 3 | La queue est bien dégagée, les pointes des fesses sont couvertes mais pas "noyées". Des boules peuvent persister mais elles sont alors parfaitement isolées. |
| NOTE 2 | La pointe des fesses apparaît nettement, sans couverture. |
| NOTE 1 | L'ensemble de la tubérosité ischiatique est perceptible. |
| NOTE 0 | On n'ignore plus rien de l'anatomie des ischions. |
| 2 - Le ligament sacro-tubéral et le détroit caudal | |
| <p>Il relie la tubérosité ischiatique (pointe de la fesse) au sacrum. Il a la particularité de se relâcher sous l'action hormonale au moment du vêlage (la vache se "casse"). Il délimite ce qu'on appelle le détroit caudal. Le maniement correspondant est le maniement des abords.</p> | |
| NOTE 5 | Le ligament est invisible, noyé. Le maniement est rendu impossible par la masse de gras. |
| NOTE 4 | Le ligament sacro-tubéral est à peine visible, (bien perceptible au maniement). Le détroit caudal reste pratiquement comblé. On prend une bonne poignée à la queue. |
| NOTE 3 | Le ligament est bien visible, couvert, d'aspect épais et arrondi. Le détroit caudal n'est plus comblé, ses limites sont planes et forment un V. On prend nettement du gras à la queue. |
| NOTE 2 | Le ligament bien isolé est légèrement couvert. La peau est souple. |
| NOTE 1 | Le ligament prend son aspect en lame, sec, et découvre un détroit caudal profond. La peau est tendue, elle adhère moins qu'en NOTE 0. |
| NOTE 0 | Le ligament en lame très sec "protège" un détroit caudal très creusé. La peau adhère sur l'arête du ligament sacro-tubéral. |
| 3 - La ligne du dos | |
| <p>La vue arrière permet de regarder les reins de l'animal et de constater si l'apophyse épineuse (voire le corps vertébral) fait saillie sur la ligne du dos.</p> <p>Remarque : Les apophyses transverses sont horizontales. Toute pente observée ne peut qu'être le fait d'un tissu sous-jacent, musculaire (faux-filet) ou graisseux.</p> | |
| NOTE 5 | Le dos est parfaitement plat ou arrondi, aucun élément osseux ne se devine. |
| NOTE 4 | On voit à peine la ligne du dos noyée dans la masse du tissu environnant. Elle est plus visible chez les animaux de sang Holstein. |
| NOTE 3 | La ligne des apophyses épineuses est nettement visible mais encore couverte, la différence de niveau entre cette ligne et le plat du rein de part et d'autre commence à se marquer. |
| NOTE 2 | La ligne des apophyses épineuses est marquée, peu couverte. |
| NOTE 1 | La ligne des apophyses épineuses est irrégulière, le corps vertébral n'est pas apparent. La différence de niveau entre l'épine dorsale et le rein est importante. |
| NOTE 0 | Le corps vertébral est apparent. |

B – QUELS CRITERES OBSERVER POUR ATTRIBUER LA NOTE "FLANC" ?

Nous allons décrire les deux critères visuels qui permettent d'attribuer la note "FLANC". Ils portent sur l'observation au niveau du flanc DROIT de :

- la pointe de la hanche,
- les apophyses transverses et épineuses.

Le dernier critère est précisé par des informations pour le maniement du travers.

Nous donnerons également des informations sur les maniements de la côte et du grasset.

1 - Pointe de la hanche.

Seul le versant antérieur de la pointe de la hanche est à regarder.

NOTE 5 La région anatomique reste repérable, mais l'épaisseur du tissu sous-jacent interdit une localisation précise de l'os (ilium).

NOTE 4 L'ilium apparaît plus nettement mais les angles restent ouverts.

NOTE 3 L'ilium fait une saillie assez nette. Il reste couvert. La crête demeure invisible.

NOTE 2 à 0 De la note 2 à 0, l'ilium est de plus en plus apparent. En NOTE 1, la crête de l'ilium devient apparente. En NOTE 0, elle est très visible.

2 - Apophyses transverses et épineuses.

La vue latérale droite permet d'estimer l'épaisseur du tissu sous-cutané bordant les apophyses transverses (correspondant au maniement du travers) et recouvrant les apophyses épineuses de la colonne vertébrale en région lombaire. La note visuelle du travers est délicate à attribuer sur le creux du flanc gauche bombé par le rumen.

NOTE 5 Aucune structure osseuse n'est repérable. Le rein est parfaitement plat et le creux du flanc (variable suivant l'état de gestation et de réplétion de la panse) est très largement comblé.

Sur la couche très importante de gras la peau roule.

NOTE 4 Le rein est encore très plat mais la colonne vertébrale est repérable. La ligne des apophyses transverses devient repérable mais la peau suit cette ligne sur une courbe très arrondie.

La peau est soulevée par la couche de gras.

NOTE 3 La ligne des apophyses épineuses est bien dessinée. Le rein n'est plus plat. La ligne des apophyses transverses est plus marquée mais l'angle n'est pas encore vif.

Le faux filet est couvert.

NOTE 2

Accentuation de la ligne des apophyses épineuses.

La ligne des apophyses transverses fait un angle vif. Les apophyses transverses commencent à s'individualiser.

La peau est souple.

NOTE 1

Accentuation des caractères précédents. On peut facilement compter les apophyses transverses en bordure.

La peau est collée.

NOTE 0

La ligne du dos est irrégulière et l'on peut presque compter les apophyses épineuses.

Les apophyses transverses sont bien individualisées jusqu'au corps vertébral.

Le faux filet est très mince, la peau est collée et très sèche.

3 - Les côtes, la dernière côte.

Les côtes sont plus ou moins couvertes et c'est un maniement sûr ; ce n'est pas un critère visuel à retenir.

On peut, sur un animal gras voir "rouler" la peau lorsqu'il se déplace. Sur un animal maigre la peau est plus "collée" aux côtes.

Le maniement s'effectue en caressant les côtes de l'avant vers l'arrière et en s'arrêtant sur la dernière côte.

NOTE 5

La peau roule considérablement sous la main. On sent un bourrelet important de gras sur la dernière côte en posant la main.

NOTE 4

En posant la main on ne sent pas la dépression du dernier espace intercostal. (La main repose à plat).

NOTE 3

La peau roule sous la main. Sans appuyer on sent bien la dépression du dernier espace intercostal.

NOTE 2

On décèle une légère pellicule de gras, la peau est souple.

NOTE 1

La peau est collée et on ne décèle aucune trace de gras.

NOTE 0

La peau est très sèche et collée.

4 - Le grasset.

NOTE 5

On saisit la graisse à pleine main.

NOTE 4

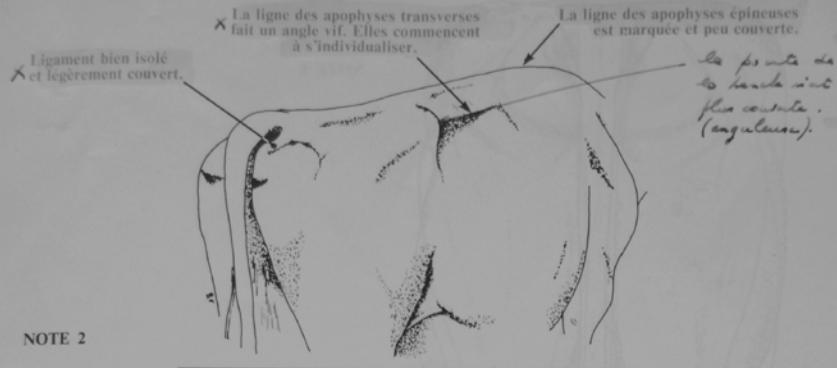
On saisit une bonne poignée de gras.

NOTE 3

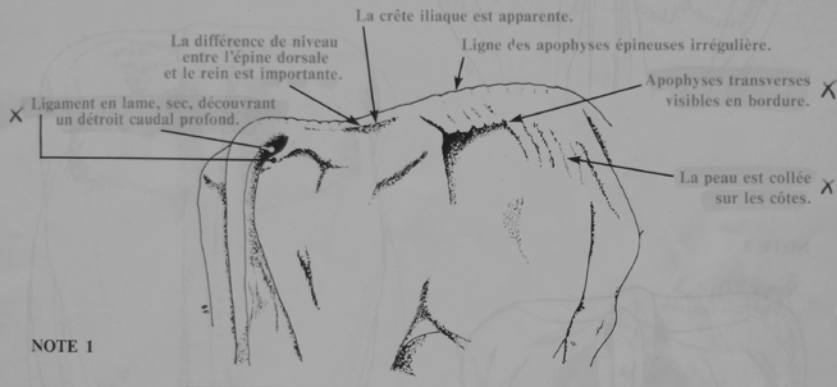
On saisit une petite poignée de graisse.

5 - Le coude, l'épaule.

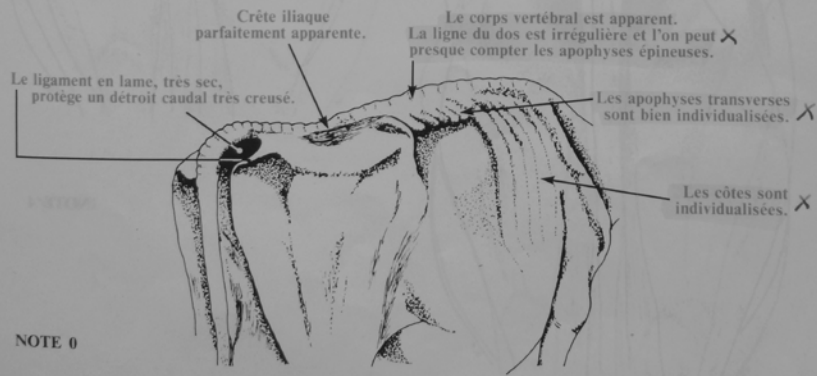
Il est impossible de donner des repères précis sur ce point, l'œil exercé fera la différence en dehors de critères explicites.



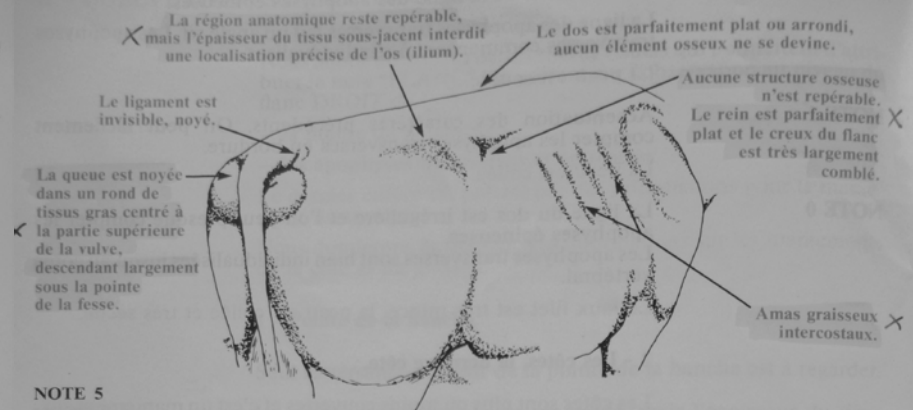
NOTE 2



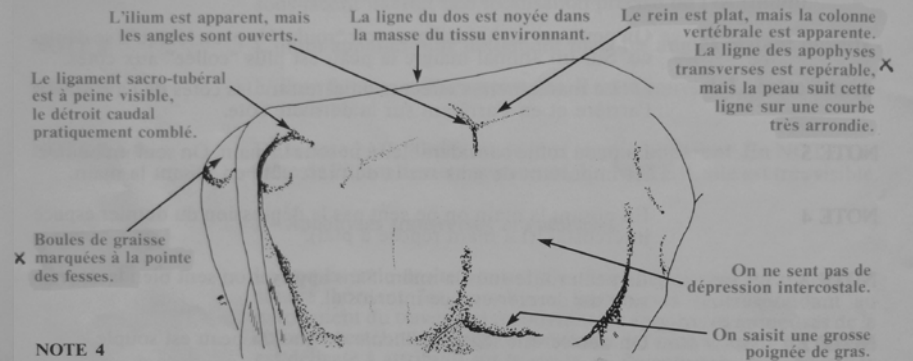
NOTE 1



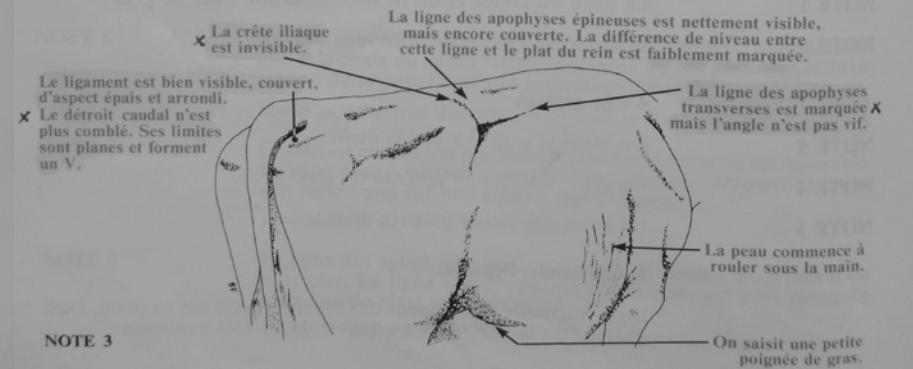
NOTE 0



NOTE 5



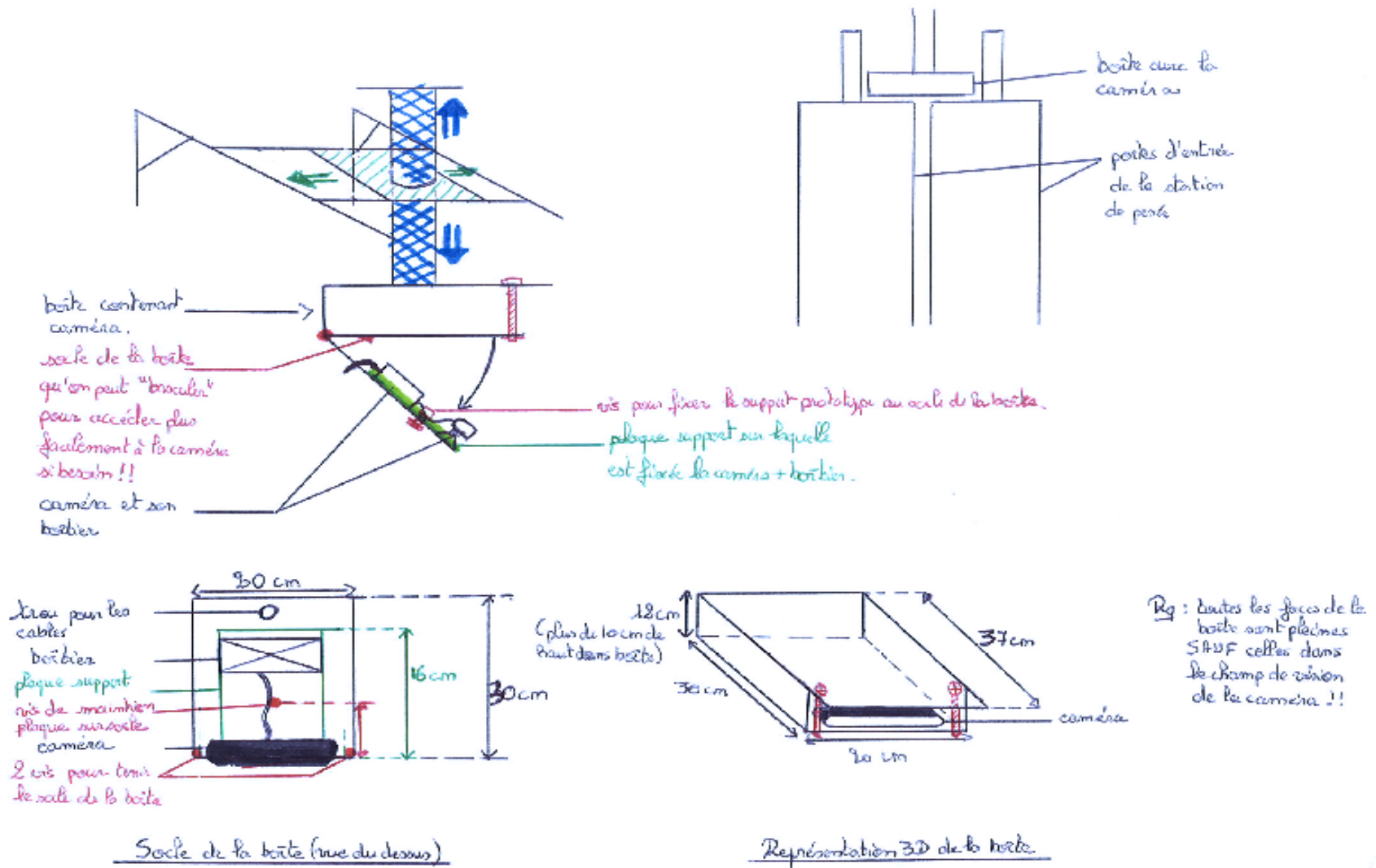
NOTE 4



NOTE 3

| <div style="text-align: right;">Note</div> <div style="text-align: left;">Critère de note Arrière</div> | 5 | 4 | 3 | 2 | 1 | 0 |
|---|--|--|--|--|--|---|
| Base de la queue Pointe des fesses | queue noyée dans un rond de tissu gras | absence de rond, masses graisseuses débordant largement la pointe des fesses | queue bien dégagée, pointe des fesses couverte mais non noyée | pointe des fesses sans couverture | ensemble de la tubérosité ischiatique perceptible | le bassin est parfaitement visible |
| Ligament sacro-tubéral | invisible noyé | à peine visible | bien visible couvert, d'aspect épais et arrondi | bien isolé, légèrement couvert | aspect en lame, sec | très sec |
| Détroit caudal | totalemement comblé | pratiquement comblé | n'est plus comblé, limites planes | profond | très profond | très creusé |
| Epine dorsale | dos plat ou arrondi | à peine visible, noyée dans le tissu environnant | nettement visible, couverte, régulière différence de niveau légère | ligne des apophyses épineuses marquée, peu couverte | ligne irrégulière, différence de niveau importante, corps vertébral non apparent | corps vertébral apparent |
| <div style="text-align: right;">Note</div> <div style="text-align: left;">Critère de note Flanc</div> | 5 | 4 | 3 | 2 | 1 | 0 |
| Pointe de la hanche | localisation précise de l'os impossible | ilium apparent angle ouvert | ilium fait saillie, reste couvert | la crête n'est pas apparente, angle bien vif | la crête devient visible | la crête est très visible |
| Apophyses transverses et épineuses | aucune structure repérable, rein plat, creux du flanc comblé | colonne vertébrale repérable, rein plat | épine dorsale bien dessinée, rein non plat, bordure des apophyses transverses nette, angle non vif | ligne des apophyses transverses fait un angle vif, on commence à pouvoir les compter | on peut compter facilement les apophyses transverses | ligne du dos très irrégulière, apophyses transverses bien individualisées |

Annexe II : Plan et localisation du système d'acquisition





Capteur Xtion

Détecteur mécanique de fermeture des portes qui déclenche

Porte d'entrée de la station de pesée



Capteur Xtion

Portes de sortie de la station

Portes d'entrée de la station

Sens de passage des vaches

Annexe III : Script des analyses statistiques réalisées avec le logiciel R : la calibration, la validation et la qualification de la méthode

```
#####
#####
#####
##### script pour la construction de la nec3d #####

## création d'une fonction pour représenter la droite des moindres rectangles:
frect <- fonction(x,y){ # droite des moindres rectangles
  b <- sqrt(var(y)/var(x))*sign(cor(x,y))
  c(mean(y)-b*mean(x),b)
}

#####
##### création des populations d'étude #####

###importation des données pour la population de calibration: ##

### création d'un tableau vide pour contenir les coordonnées pour la méthode "3points" population calibration et
### validdiff:
points<-as.matrix(read.table("F:/resultats_nec3d/coordonneesacp/repertoire.txt",sep=";"))
points ## contient les chemins d'accès aux fichiers contenant les surfaces 3D
nom<-as.matrix(read.table("F:/resultats_nec3d/coordonneesacp/repertoire.txt",sep="/"))
nom ## contient le nom des fichiers utilisés
p<-nrow(points)
coordonnees<-matrix(0,nrow=716,ncol=57) ## création d'une matrice de zéro
description<-matrix(0,nrow=716,ncol=6) ## création d'une matrice de zéro

k<-0
for (i in 1:p)
{env<-points[i] ## prend le chemin d'accès de la ligne "i" de le base de données points
  troispnts<-read.table(paste(env,"/coordonnees.txt",sep=""),sep=";")
  ## charge les coordonnées dans le repère de l'ACP
  ## la population associée au chemin d'accès "env" tel
  ## que le séparateur soit ";"
  vache<-as.matrix(read.table(paste(env,"/coordonnees.txt",sep=""),sep="_")[,2])
  ## extrait la 2ème colonne contenant le numéro des vaches
  test<-as.matrix(read.table(paste(env,"/coordonnees.txt",sep=""),sep="_")[,1])
  ## extrait la nature de la population: calibration, validation ou reproductibilité
```

```

colnames(vache)=c("vache") ## attribue "vache" comme titre à la colonne de la base de données "vache"
for (j in 1:nrow(vache))
{point1<- as.numeric(scan(paste(env,as.character(troispoints$V1)[j],sep="/"), what="",dec=".",sep=","))
  ## importe les coordonnées de la vache de la jème ligne du fichier troispoints
  point1<-t(as.matrix(point1)) ## transpose la matrice qui était sous forme de colonne avec les
  ## coordonnées en lignes alors qu'on veut la fusionner avec "coordonnees" qui contient les
  ## coordonnées en colonnes.
  k<-k+1 ## compteur de la boucle
  description[k,1]<-nom[i,7] ## ajoute la nature de la méthode à la base de données: "4points" ou "3points"
  description[k,2]<-nom[i,8] ## ajoute la nature de la résolution à la base de données: "150150" ou
  ## "100100" ou "5050"
  description[k,3]<-nom[i,9] ## ajoute la nature de la population à la base de données: "calibration" ou
  ## "validation" ou "image" ou "pointage"
  description[k,4]<-vache[j] ## ajoute l'identifiant de la vache à la base de données
  if (description[k,3]=="pointage" || description[k,3]=="image"){description[k,5]<-
trunc(as.numeric(vache[j])/1000)}else{description[k,5]<- "NA"}
  ## ajoute si c'est une vache de la population de calibration "CowDB" ou de validation de
  ## valididem "Test"
  description[k,6] <-test[j] ## ajoute le numéro de la série de l'essai de reproductibilité à la matrice
  ## "coordonnées": 1 à 8
  coordonnees[k,]<-point1}} ## ajoute les coordonnées de la vache dans la ligne k

coordonnees<-as.data.frame(coordonnees) ## transforme la matrice "coordonnées" en base de données
vache<-as.character(vache) ## transforme la variable "vache" en variable qualitative
descripteur<-as.data.frame(description)
colnames(descripteur)=c("methode","grille","population","vache","serie","pop")
## identifie chacune des colonnes en leur donnant un titre
necr<-read.table("C:/Documents and Settings/afischer/Mes
documents/program3d/ReleaseCow3D/Cows/nec.csv",sep=";",header=TRUE,dec=",")
## importation du fichier contenant les NEC des vaches

points4<-cbind(descripteur,coordonnees) ## accole les 2 bases de données par colonne
quatreponts<-merge(points4,necr,by=c("vache")) ## fusionne les 2 bases de données selon la variable "vache"
## commune aux 2 bases de données

dim(quatrepoints)
dim(descripteur)
quatreponts
colnames(quatrepoints)<-gsub("V","cp",colnames(quatrepoints),fixed=TRUE)
write.table(quatrepoints,file=paste("F:/resultats_nec3d/calibration/","quatreponts.csv",sep=""),sep=";",dec="," ,row.names=FALSE)
## exporte la base de données "quatreponts" au chemin d'accès renseigné

```

```
#####
## création d'une table contenant les 4 modèles:
listemodele<-as.data.frame(matrix(nrow=56,ncol=0))

## création d'une table qui contiendra tous les résidus de tous les modèles calculés dans la suite pour les
## 2 populations de validation:
idem<-subset(quatreponts,quatreponts$methode=="4points3remplissage" & quatreponts$grille=="150150" &
quatreponts$population=="valididem")
    ## extrait la un partie de la base de données "quatreponts" satisfaisant les conditions
    ## renseignées entre
    ## parenthèse: "valididem"
diff<-subset(quatreponts,quatreponts$methode=="4points3remplissage" & quatreponts$grille=="150150" &
quatreponts$population=="validdiff" & quatreponts$pop=="cowTest")
    ## extrait la population "validdiff" de la base quatreponts
resdiff<-as.data.frame(matrix(0,nrow=25,ncol=1))
    ## création de la base de données qui contiendra les résidus des modèles pour la population
    ## validdiff
resdiff<-as.data.frame(diff$vache)
residem<-as.data.frame(matrix(0,nrow=25,ncol=1))
    ## création de la base de données qui contiendra les résidus des modèles pour la population
    ## valididem
residem<-as.data.frame(idem$vache)

quatreponts$vache<-as.factor(quatreponts$vache)
quatreponts$serie<-as.factor(quatreponts$serie)
quatreponts$grille<-as.factor(quatreponts$grille)
summary(quatreponts)
```

```

#####
#####
##### méthode 4 points imitant 3 points pour une grille 150x150 #####

## extraction des sous-population "calibration" (construction), "nec3d_idem" et "nec3d_diff" (pour "valididem"
## et "validdiff"), "image" et "pointage" (pour l'analyse de la reproductibilité et de la répétabilité):

construction<-subset(quatreponts,quatreponts$methode=="4points3remplissage" & quatreponts$grille=="150150" &
quatreponts$pop=="cowDB" )
pointage<-subset(quatreponts,quatreponts$methode=="4points3remplissage" & quatreponts$grille=="150150" &
quatreponts$population=="pointage" )
image<-subset(quatreponts,quatreponts$methode=="4points3remplissage" & quatreponts$grille=="150150" &
quatreponts$population=="image" )
nec3d_idem<-subset(quatreponts,quatreponts$methode=="4points3remplissage" & quatreponts$grille=="150150" &
quatreponts$population=="valididem" )
nec3d_diff<-subset(quatreponts,quatreponts$methode=="4points3remplissage" & quatreponts$grille=="150150" &
quatreponts$population=="validdiff" & quatreponts$pop=="cowTest" )

summary(construction)  ## résume la base "construction", pour s'assurer que chacune des variables est bien définie
                      ## (qualitative vs. quantitative)

##sélection du meilleur modèle cherchant à expliquer les variations observées de la NEC en fonction des coordonnées des
## surfaces 3D dans la population de calibration dans le repère de l'ACP:

adjrcarre<-
data.frame(modele=0,étape=0,rcarré=0,rcarrévalididem=0,rcarrévaliddiff=0,rcarrépoints=0,rcarréimage=0,erreur=0,erreurva
lididem=0,erreurvaliddiff=0,erreurpoints=0,erreurimage=0)
      ## création d'une table contenant pour chaque étape du stepwise le r2 ajusté du modèle et le r2
      ## ajusté du modèle sur la population validation externe (rcarrévalididem et rcarrévaliddiff),
      ## sur les populations pour la répétabilité (rcarrépoints et rcarréimage) et l'étape du stepwise.
      ## Cette étape renseigne que pour démarrer le modèle ne contient aucune variable, adjr2=0

for (i in 1:57)
  {stepi<-step(lm(nec~1,construction),nec ~
cp1+cp2+cp3+cp4+cp5+cp6+cp7+cp8+cp9+cp10+cp11+cp12+cp13+cp14+cp15+cp16+cp17+cp18+cp19+cp20+cp21+cp22+cp23+cp24+cp25+cp2
6+cp27+cp28+cp29+cp30+cp31+cp32+cp33+cp34+cp35+cp36+cp37+cp38+cp39+cp40+cp41+cp42+cp43+cp44+cp45+cp46+cp47+cp48+cp49+cp
50+cp51+cp52+cp53+cp54+cp55+cp56+cp57,steps=i,direction="both",data=construction)
      ## sélection "pas à pas" du meilleur modèle de régression linéaire, on part du modèle nul ne
      ## contenant aucune variable, la sélection ajoute ou supprime une variable à chaque étape parmi

```

```

    ## les 57 variables.
outputi<-summary(stepi)  ## résume la sélection "pas à pas"
adjrsqi<-outputi$adj.r.squared  ## extraction du R2 ajusté
erreur1<-sd(outputi$residuals) ## extraction de l'erreur résiduelle du modèle
## population de validation avec les individus identiques à la population de construction: ##
prediction_idem<-predict(stepi,nec3d_idem,se.fit=TRUE)  ## calcul les estimations de NEC sur la
## population "valididem"

predict3didem<-as.data.frame(prediction_idem$fit)
stepvalid_idem<-cbind(nec3d_idem,predict3didem)  ## accole les NEC mesurées aux prédictions
valididem.stepi<-lm(nec~prediction_idem$fit,data=stepvalid_idem)
## régression linéaire simple de la NEC (mesurée) en fonction de la nec3d (NEC estimée)
adjrsqvalididem<-summary(valididem.stepi)$adj.r.squared  ## extraction du R2 ajusté
erreurvalidid<-sd(summary(valididem.stepi)$residuals) ## extraction de l'erreur résiduelle
## population de validation avec les individus différents de la population de construction: ##
prediction_diff<-predict(stepi,nec3d_diff,se.fit=TRUE)  ## calcul les estimations de NEC sur la
## population "validdiff"

predict3ddiff<-as.data.frame(prediction_diff$fit)
stepvalid_diff<-cbind(nec3d_diff,predict3ddiff)
validdiff.stepi<-lm(nec~prediction_diff$fit,data=stepvalid_diff)
adjrsqvaliddiff<-summary(validdiff.stepi)$adj.r.squared
erreurvaliddif<-sd(summary(validdiff.stepi)$residuals)
## population pour tester la répétabilité du pointage: ##
prediction_pts<-predict(stepi,pointage,se.fit=TRUE)  ## calcul les estimations de NEC sur la
## population "pointage"

predict3dpts<-as.data.frame(prediction_pts$fit)
steppts<-cbind(pointage,predict3dpts)
pts.stepi<-lm(nec~prediction_pts$fit,data=steppts)
adjrsqpts<-summary(pts.stepi)$adj.r.squared
erreurpts<-sd(summary(pts.stepi)$residuals)
## population pour tester la répétabilité de l'indicateur sur un même individu pour des images
## différentes:
prediction_im<-predict(stepi,image,se.fit=TRUE)  ##calcul les estimations de NEC sur la population "image"
predict3dim<-as.data.frame(prediction_im$fit)
stepim<-cbind(image,predict3dim)
im.stepi<-lm(nec~prediction_im$fit,data=stepim)
adjrsqim<-summary(im.stepi)$adj.r.squared
erreurim<-sd(summary(im.stepi)$residuals)
## construction d'une base de données contenant tous les R2 et les erreurs de tous les modèles pour
## chacune des populations:
adjrcarrei<-
data.frame(modele=0,étape=i,rcarré=adjrsqi,rcarrévalididem=adjrsqvalididem,rcarrévaliddiff=adjrsqvaliddiff,rcarrépoints
=adjrsqpts,rcarréimage=adjrsqim,erreur=erreur1,erreurvalididem=erreurvalidid,erreurvaliddiff=erreurvaliddif,erreurpoint
s=erreurpts,erreurimage=erreurim)

```



```

adjrcarre<-rbind(adjrcarre,adjrcarrei) ## fusionne l'ancienne table "adjrcarre" avec la nouvelle ligne
## calculée pour l'étape i
}

## représentation graphique de l'évolution de la qualité des modèles en fonction du nombre de composantes
## dans le modèle, R2=f(i): ##
modelebest<-stepi$call ## extrait le modèle complet de la sélection "pas à pas"
modelebest
as.data.frame(stepi$call)
write.table(stepi$coefficients,file=paste("F:/resultats_nec3d/calibration/",paste(construction$methode[1],construction$
grille[1],construction$population[1],construction$pop[1],sep=""),"meilleurmodele.csv",sep=""),sep=";",dec=",")
## extraction du nom des variables du modèle, pour pouvoir libeller l'axe des abscisses de nos graphiques:
modell<-gsub("nec ~ ","",modelebest, fixed=TRUE) ## remplace exactement les caractères "nec ~ " du modèle
## par un vide "" avec l'option fixed=TRUE qui permet de
## remplacer exactement les caractères

modell
model<-gsub("+","",modell[2],fixed=TRUE) ## remplace le caractère "+" par ""
model
composante<-strsplit(model," ") ## coupe la chaîne de caractères au niveau des " "
composante<-noquote(composante[[1]]) ## retire les "" de la chaîne de caractères
composante
composante<-as.matrix(as.factor(composante)) ## transform la variable "composante" en variable qualitative
## et la chaîne de caractères en matrice

composante
dim(composante)
cp<-as.data.frame((as.matrix(composante))) ## transformation en tableau de données
colnames(cp)=paste(construction$methode[1],construction$grille[1],construction$population[1],construction$pop[1],sep=" "
)
dim(cp) ## donne le nombre de cp du meilleur modèle
adjrcarre
listemodele<-cbind(listemodele,cp)

## exporte la table "adjrcarre" contenant l'évolution de la qualité des modèles du stepwise:
write.table(adjrcarre,file=paste("F:/resultats_nec3d/calibration/",paste(construction$methode[1],construction$grille[1]
),sep=""),sep=";",dec=",",append=FALSE,quote=TRUE)

### traçons les graphes pour voir l'évolution des modèles au fûr et à mesure de l'avancement du stepwise:
windows() ##### ouvre la console affichant les graphes sous R, cette étape est indispensable si on veut
##### exporter automatiquement les graphes sinon on peut les enregistrer manuellement dans la
##### fenêtre des graphes.

```

```

barplot(adjrcarre$rcarré,names.arg=adjrcarre$étape,main="évolution de la qualité des modèles du stepwise",xlab="nombre
d'étapes dans la sélection par stepwise",ylab="R-carré ajusté",ylim=c(0,1))
  ## représente graphiquement l'évolution du R2ajusté des étapes du stepwise sous forme d'histogramme
savePlot(filename=paste("F:/resultats_nec3d/calibration/",paste(construction$methode[1],construction$grille[1],construc
tion$population[1],construction$pop[1]),sep=""),type="jpeg",device=dev.cur())
  ## enregistre le graphe sous format .jpeg sur mon disque dur externe (F:). La commande path.expand
renseigne
  ## le chemin d'accès où R doit enregistrer le document!

## représentation graphique de l'évolution du R2ajusté des modèles de prédiction sur les populations
## "valididem", "validdiff", "pointage" et "image":
barplot(adjrcarre$rcarrévalididem,names.arg=adjrcarre$étape,main=paste("évolution de la qualité des modèles du
stepwise","appliqués sur la population de validation","avec les vaches identiques à la population de
construction",sep="\n"),xlab="nombre d'étapes dans la sélection par stepwise",ylab="R-carré ajusté",ylim=c(0,1))
savePlot(filename=paste("F:/resultats_nec3d/calibration/",paste(nec3d_idem$methode[1],nec3d_idem$grille[1],nec3d_idem$popu
lation[1]),sep=""),type="jpeg",device=dev.cur())
barplot(adjrcarre$rcarrévaliddiff,names.arg=adjrcarre$étape,main=paste("évolution de la qualité des modèles du
stepwise","appliqués sur la population de validation", "avec les vaches différentes de la population de
construction",sep="\n"), xlab="nombre d'étapes dans la sélection par stepwise",ylab="R-carré ajusté",ylim=c(0,1))
savePlot(filename=paste("F:/resultats_nec3d/calibration/",paste(nec3d_diff$methode[1],nec3d_diff$grille[1],nec3d_diff$popu
lation[1],nec3d_diff$pop[1],sep=""),sep=""),type="jpeg",device=dev.cur())
barplot(adjrcarre$rcarrépoints,names.arg=adjrcarre$étape,main=paste("évolution de la qualité des modèles du
stepwise","appliqués sur la population test pour la répétabilité du pointage",sep="\n"),xlab="nombre d'étapes dans la
sélection par stepwise",ylab="R-carré ajusté",ylim=c(0,1))
savePlot(filename=paste("F:/resultats_nec3d/calibration/",paste(pointage$methode[1],pointage$grille[1],pointage$populat
ion[1],sep=""),sep=""),type="jpeg",device=dev.cur())
barplot(adjrcarre$rcarréimage,names.arg=adjrcarre$étape,main=paste("évolution de la qualité des modèles du
stepwise","appliqués sur la population test pour la répétabilité intra-individu",sep="\n"),xlab="nombre d'étapes dans
la sélection par stepwise",ylab="R-carré ajusté",ylim=c(0,1))
savePlot(filename=paste("F:/resultats_nec3d/calibration/",paste(image$methode[1],image$grille[1],image$population[1]),s
ep=""),type="jpeg",device=dev.cur())

## de la même manière, représentation graphique de l'évolution des erreurs résiduelles du modèle:
barplot(adjrcarre$erreur,names.arg=adjrcarre$étape,main="évolution de la qualité des modèles du stepwise",xlab="nombre
d'étapes dans la sélection par stepwise",ylab="erreur de prediction",ylim=c(0,0.8))
savePlot(filename=paste("F:/resultats_nec3d/calibration/","erreur",paste(colnames(adjrcarre[,8]),construction$methode[1
],construction$grille[1],construction$population[1],construction$pop[1]),sep=""),type="jpeg",device=dev.cur())

barplot(adjrcarre$erreurvalididem,names.arg=adjrcarre$étape,main=paste("évolution de la qualité des modèles du
stepwise","appliqués sur la population de validation","avec les vaches identiques à la population de
construction",sep="\n"),xlab="nombre d'étapes dans la sélection par stepwise",ylab="erreur de
prediction",ylim=c(0,0.8))

```

```
savePlot(filename=paste("F:/resultats_nec3d/calibration/", "erreur", paste(colnames(adjrcarre[,9]), nec3d_idem$methode[1], nec3d_idem$grille[1], nec3d_idem$population[1]), sep=""), type="jpeg", device=dev.cur())
```

```
barplot(adjrcarre$erreurvaliddiff, names.arg=adjrcarre$étape, main=paste("évolution de la qualité des modèles du stepwise", "appliqués sur la population de validation", "avec les vaches différentes de la population de construction", sep="\n"), xlab="nombre d'étapes dans la sélection par stepwise", ylab="erreur de prediction", ylim=c(0,0.8))
```

```
savePlot(filename=paste("F:/resultats_nec3d/calibration/", "erreur", paste(colnames(adjrcarre[,10]), nec3d_diff$methode[1], nec3d_diff$grille[1], nec3d_diff$population[1], nec3d_diff$pop[1]), sep=""), type="jpeg", device=dev.cur())
```

```
barplot(adjrcarre$erreurpoints, names.arg=adjrcarre$étape, main=paste("évolution de la qualité des modèles du stepwise", "appliqués sur la population test pour la répétabilité du pointage", sep="\n"), xlab="nombre d'étapes dans la sélection par stepwise", ylab="erreur de prediction", ylim=c(0,0.8))  
savePlot(filename=paste("F:/resultats_nec3d/calibration/", "erreur", paste(colnames(adjrcarre[,11]), pointage$methode[1], pointage$grille[1], pointage$population[1]), sep=""), type="jpeg", device=dev.cur())
```

```
barplot(adjrcarre$erreurimage, names.arg=adjrcarre$étape, main=paste("évolution de la qualité des modèles du stepwise", "appliqués sur la population test pour la répétabilité intra-individu", sep="\n"), xlab="nombre d'étapes dans la sélection par stepwise", ylab="erreur de prediction", ylim=c(0,0.8))  
savePlot(filename=paste("F:/resultats_nec3d/calibration/", "erreur", paste(colnames(adjrcarre[,12]), image$methode[1], image$grille[1], image$population[1]), sep=""), type="jpeg", device=dev.cur())
```

```
## représente graphiquement les valeurs de NEC estimées par le modèle en fonction des NEC mesurées:
```

```
windows()  
plot(stepi$fitted.values, construction$nec, type="p", main=paste("qualité du modèle de calibration", paste(construction$methode[1], construction$grille[1], sep=""), sep=""), xlab="nec prédite", ylab="nec observées", xlim=c(0,5), ylim=c(0,5))  
abline(lm(construction$nec~stepi$fitted.values))  
abline(a=0,b=1,col="red",lty=2) ## représentation graphique de la lère bissectrice, sur le même graphe  
savePlot(filename=paste("F:/resultats_nec3d/calibration/", "modelecalib", paste(construction$methode[1], construction$grille[1], sep=""), sep=""), type="jpeg")  
## exporte le graphe
```

```
#####  
##### processus de caractérisation du modèle: répétabilité et reproductibilité #####
```

```
## vérifions que les tableaux de données sont conformes: ##
```

```
summary(pointage)  
rownames(pointage)<-pointage[,1]  
rownames(image)<-image[,1]
```

```

summary(image)

#####
## calcul de la répétabilité de l'identification des points: ##

## pour cela, il faut prédire la nec3d à partir du modèle retenu à l'issu de la stepwise, modelebest
prediction_pts<-predict(stepi,pointage,se.fit=TRUE)  ## estimation des nec3d de la population à partir
                                                    ## du modèle retenu
valeurspredites_pts<-prediction_pts$fit  ## extraction des valeurs estimées
valeurspredites_pts
## les sorties sont sous forme d'un vecteur, transformons le en tableau et fusionnons le avec la base
## de données pointage:
nec3dpts<-as.data.frame(valeurspredites_pts)
test_pointage<-cbind(pointage,nec3dpts)
test_pointage

## création d'une matrice vide qui contiendra les critères calculés pour la répétabilité et la
## reproductibilité:
fidelite<-matrix("",nrow=2,ncol=6)  ## créer une matrice vide
fidelite[1,1]<-"répétabilité"  ## identifie la lère ligne aux résultats de répétabilité
fidelite[2,1]<-"reproductibilité"  ## identifie la lère ligne aux résultats de reproductibilité

## les identifiants des vaches utilisés ne permettent pas de dissocier l'effet série de l'effet vache,
## ils sont confondus,on supprime le premier chiffre de l'identifiant:
id<-substring(pointage$vache,2,4)  ## supprime le premier chiffre en conservant les 3 derniers caractères
                                   ## de l'identifiant: c'est-à-dire les caractères en position 2 à 4 dans le
                                   ## nombre, par exemple pour 1075 elle conserve 075.

id
id<-gsub("575","075",id,fixed=TRUE)  ## pour éviter que 2 vaches 575 soient dans la table pointage, on
                                   ## attribue un identifiant différent

id<-as.factor(id)
id
test_pointage<-cbind(test_pointage,id)  ## ce nouvel identifiant est ajouté à la table "pointage" mais
                                       ## n'écrase pas l'identifiant initial "vache"

test_pointage

## réalisation de l'anova pour analyser l'erreur liée au pointage, c'est-à-dire en retirant l'effet vache,
## "id":
anova_pts<-aov(valeurspredites_pts~id,data=test_pointage)
summary(anova_pts)
dllr=48-anova_pts$rank+1
sr2=var(anova_pts$residuals)*47/dllr  ## calcul de la variance de répétabilité
sr2

```

```

M=mean(valeurspredites_pts) ## moyenne des nec3d
r=sqrt(sr2) ## écart-type de répétabilité
r
## pour les mêmes conditions expérimentales,même ferme, même personne, même méthode, la différence
## maximale au risque de ## 5% qui sépare 2 résultats sur une même acquisition vaut ##
r ## unité de NEC sur un échantillon dont la moyenne est proche de ##
M ##.#
cvr=r/M*100 ## coefficient de variation de répétabilité ##
cvr

fidelite[1,2]<-dllr
fidelite[1,3]<-sr2
fidelite[1,4]<-r
fidelite[1,5]<-M
fidelite[1,6]<-cvr

#####
## calcul de la reproductibilité intra-ferme de l'indicateur:##
prediction_im<-predict(stepi,image,se.fit=TRUE) ## estimation des nec3d de la population à partir du
## modèle conservé

valeurspredites_im<-prediction_im$fit ## extraction des estimations
valeurspredites_im
#les sorties sont sous forme d'un vecteur, transformons le en tableau et fusionnons le avec la base de
## données pointage:
nec3dim<-as.data.frame(valeurspredites_im)
test_image<-cbind(image,nec3dim)
test_image

## les identifiants des vaches utilisés ne permettent pas de dissocier l'effet série de l'effet vache,
## ils sont confondus,on supprime le premier chiffre de l'identifiant:
idr<-substring(image$vache,2,4) ## supprime le premier chiffre en conservant les 3 derniers caractères de
l'identifiant:
## c'est-à-dire les caractères en position 2 à 4 dans le nombre,
## par exemple pour 1075 elle conserve 075.

idr
idr<-gsub("575","075",idr,fixed=TRUE)
idr<-as.factor(idr)
idr
test_image<-cbind(test_image,idr)
test_image

## réalisation de l'anova pour analyser l'erreur de reproductibilité, c'est-à-dire en retirant l'effet
## vache, "idr":

```

```

anova_im<-aov(valeurspredites_im~idr,data=test_image)
summary(anova_im) dllrim=48-
anova_im$rank+1
srim2=var(anova_im$residuals)*47/dllrim
srim2
Mim=mean(valeurspredites_im)
rim=sqrt(srim2)
rim

## pour les mêmes conditions expérimentales,même ferme, même personne, même méthode, la différence
## maximale au risque de ## 5% qui sépare 2 résultats sur une même acquisition vaut ##
rim
# unité de NEC sur un échantillon dont la moyenne est proche de #
Mim #.#
cvrim=rim/Mim*100 ##coefficient de variation de répétabilité ##
cvrim
fidelite[2,2]<-dllrim
fidelite[2,3]<-srim2
fidelite[2,4]<-rim
fidelite[2,5]<-Mim
fidelite[2,6]<-cvrim

fidelite<-as.data.frame(fidelite)
colnames(fidelite)<-c("fidélité","dll","srcarré","répétabilité ou reproductibilité","moyenne","CV") ## libelé des
colonnes de "fidelite"
write.table(fidelite,file=paste("F:/resultats_nec3d/fidelite/",paste(pointage$methode[1],pointage$grille[1]),sep=""),se
p=";",dec=" ",row.names=FALSE)
## exporte la table "fidelite"
#####
##### processus de validation #####

## vérifions les bases de données des populations de validation: ##

summary(nec3d_idem)
rownames(nec3d_idem)<-nec3d_idem[,1]
summary(nec3d_diff)
rownames(nec3d_diff)<-nec3d_diff[,1]

## validation sur la population de validation contenant les mêmes vaches que la population de construction,
"valididem":
prediction_idem<-predict(stepi,nec3d_idem,se.fit=TRUE) ## calcul des nec3d prédites par le meilleur modèle
nec3didem<-prediction_idem$fit ## extraction des estimations, nec3d

```

```

nec3didem
nec3did<-as.data.frame(nec3didem)
valid_idem<-cbind(nec3d_idem,nec3did)
valid_idem

## qualité de l'estimation pour des vaches identiques à la population de construction mais avec des
## acquisitions différentes:
modele_valid_idem<-lm(nec~nec3didem,data=valid_idem) ## régression linéaire de NEC mesurée en fonction
## de NEC estimée
rcarre_valididem<-summary(modele_valid_idem)$adj.r.squared ## extraction du R2ajusté
varianceerreur_valididem<-summary(modele_valid_idem)$sigma ## extraction de l'erreur résiduelle

## création d'une matrice qui contiendra les résultats de validation pour les 2 populations:
validdiff<-matrix("",nrow=2,ncol=7) ## création d'une matrice vide
validdiff[1,1]<-"valididem" ## identifie la 1ère ligne comme les résultats de validation pour la
## population "valididem"
validdiff[2,1]<-"validdiff" ## identifie la 1ère ligne comme les résultats de validation pour la
## population "validdiff"

## calcul des différents critères pour la validation:
biais2_3id<-(mean(valid_idem$nec)-mean(valid_idem$nec3didem))^2 ## calcul du biais
NU3id<-(1-modele_valid_idem$coefficients[2])^2*var(valid_idem$nec3didem) ## calcul de l'erreur liée à
## la pente
LC3id<-(1-cor(valid_idem$nec,valid_idem$nec3didem)^2)*var(valid_idem$nec) ## calcul de l'erreur
## inexpliquée
MSE3id<-biais2_3id+NU3id+LC3id ## calcul de l'erreur globale
effetbiais_3id<-biais2_3id/MSE3id ## calcul du biais en % de l'erreur globale
effetpente_3id<-NU3id/MSE3id ## calcul de l'erreur liée à la pente en % de l'erreur globale
effetresidu_3id<-LC3id/MSE3id ## calcul de l'erreur inexpliquée en % de l'erreur globale

## ajout des résultats de validation dans la table "validdiff":
validdiff[1,2]<-MSE3id
validdiff[1,3]<-effetbiais_3id
validdiff[1,4]<-effetpente_3id
validdiff[1,5]<-effetresidu_3id
validdiff[1,6]<-summary(modele_valid_idem)$r.squared
validdiff[1,7]<-sd(summary(modele_valid_idem)$residuals)

##### validation sur la population de validation contenant des vaches différents de la
## population de construction:

```

```

## on procède de la même manière que pour la validation sur la population "validdiff":
prediction_diff<-predict(stepi,nec3d_diff,se.fit=TRUE)
nec3ddiff<-prediction_diff$fit
nec3ddiff
nec3ddi<-as.data.frame(nec3ddiff)
valid_diff<-cbind(nec3d_diff,nec3ddi)
valid_diff

#####
## qualité de l'estimation pour des vaches autres que celles de la population de construction mais
## avec des acquisitions ## différentes: ##

modele_valid_diff<-lm(nec~nec3ddiff,data=valid_diff)
rcarre_validdiff<-summary(modele_valid_diff)$adj.r.squared
rcarre_validdiff
varianceerreur_validdiff<-summary(modele_valid_diff)$sigma
varianceerreur_validdiff

biais2_3dif<-(mean(valid_diff$nec)-mean(valid_diff$nec3ddiff))^2
NU3dif<-(1-modele_valid_diff$coefficients[2])^2*var(valid_diff$nec3ddiff)
LC3dif<-(1-cor(valid_diff$nec,valid_diff$nec3ddiff)^2)*var(valid_diff$nec)
MSE3dif<-biais2_3dif+NU3dif+LC3dif
effetbiais_3dif<-biais2_3dif/MSE3dif
effetpente_3dif<-NU3dif/MSE3dif
effetresidu_3dif<-LC3dif/MSE3dif
validdiff[2,2]<-MSE3dif
validdiff[2,3]<-effetbiais_3dif
validdiff[2,4]<-effetpente_3dif
validdiff[2,5]<-effetresidu_3dif
validdiff[2,6]<-summary(modele_valid_diff)$r.squared
validdiff[2,7]<-sd(summary(modele_valid_diff)$residuals)

validdiff<-as.data.frame(validdiff)
colnames(validdiff)<-c("validation","biaisglobal","biaismoyenne","biaispente","biaisinexpliqué","rcarré","erreur
prediction")
    ## identifie les colonnes de la table "validdiff"
write.table(validdiff,file=paste("F:/resultats_nec3d/validation/",paste(valid_diff$methode[1],valid_diff$grille[1],sep=
""),sep=""),sep=";",dec="," ,row.names=FALSE)
    ## exporte la table sur mon disque dur

## représentation graphique des nec3d en fonction de la nec pour les 2 populations de validation:
windows()

```



```

plot(valid_diff$nec3ddiff,valid_diff$nec,type="p",main="validation de nec3d sur la population
validdiff",col=4,xlim=c(0,5),ylim=c(0,5))
abline(modele_valid_diff,col=4) ## représentation graphique du modèle
abline(a=0,b=1,lty=2,col=1) ## représentation graphique de la première bissectrice
savePlot(filename=paste("F:/resultats_nec3d/validation/",paste(valid_diff$methode[1],valid_diff$grille[1],valid_diff$po
pulation[1],sep=""),sep=""),type="jpeg",device=dev.cur())

```

```

plot(valid_idem$nec3didem,valid_idem$nec,type="p",main="validation de nec3d sur la population
valididem",col=2,xlim=c(0,5),ylim=c(0,5))
abline(modele_valid_idem,col=2) ## représentation graphique du modèle
abline(a=0,b=1,lty=2,col=1) ## représentation graphique de la première bissectrice
savePlot(filename=paste("F:/resultats_nec3d/validation/",paste(valid_idem$methode[1],valid_idem$grille[1],valid_idem$po
pulation[1],sep=""),sep=""),type="jpeg",device=dev.cur())

```

```

#####
##### extraction des résidus des modèles appliqués sur les 2 populations de validation #####

```

```

## extraction des résidus pour chacune des populations de validation:

```

```

residus_validdiff<-as.data.frame(modele_valid_diff$residuals) ## extrait les résidus du modèle de validation
## pour la population "validdiff"
residus_validdiff<-as.data.frame(residus_validdiff[,1])
colnames(residus_validdiff)<-paste(valid_diff$methode[1],valid_diff$grille[1],sep="")
## identifie les colonnes de la table "residus_validdiff"

```

```

residus_valididem<-as.data.frame(modele_valid_idem$residuals)
## extrait les résidus du modèle de validation pour la population "valididem"
residus_valididem<-as.data.frame(residus_valididem[,1])
colnames(residus_valididem)<-paste(valid_idem$methode[1],valid_idem$grille[1],sep="")
## identifie les colonnes de la table "residus_valididem"

```

```

resdiff<-cbind(resdiff,residus_validdiff) ## création d'une table contenant les résidus pour la population "validdiff"
write.table(resdiff,file=paste("F:/resultats_nec3d/comparaison/","residusdiff.csv",sep=""),sep=";",dec=" ",row.names=FA
LSE)

```

```
## exporte la table sur le disque dur
residem<-cbind(residem,residus_valididem) ## création d'une table contenant les résidus pour la population "valididem"
write.table(residem,file=paste("F:/resultats_nec3d/comparaison/", "residusidem.csv", sep=""), sep=";", dec="," ,row.names=FALSE)
## exporte la table sur le disque dur
```

```
#####
##cette procédure entière est répétée trois fois: pour la méthode "3points" avec la grille 100x100;
## "3points" avec la grille 50x50 et la méthode "4points":
```

```
##### DEBUT DE REPETITION DE LA PROCEDURE #####
```

```
#####
#####
##### méthode 4 points imitant 3 points pour une grille 100x100 #####
```

```
construction<-subset( quatrepoints, quatrepoints$methode=="4points3remplissage" & quatrepoints$grille=="100100" &
quatrepoints$pop=="cowDB" )
pointage<-subset( quatrepoints, quatrepoints$methode=="4points3remplissage" & quatrepoints$grille=="100100" &
quatrepoints$population=="pointage" )
image<-subset( quatrepoints, quatrepoints$methode=="4points3remplissage" & quatrepoints$grille=="100100" &
quatrepoints$population=="image" )
nec3d_idem<-subset( quatrepoints, quatrepoints$methode=="4points3remplissage" & quatrepoints$grille=="100100" &
quatrepoints$population=="valididem" )
nec3d_diff<-subset( quatrepoints, quatrepoints$methode=="4points3remplissage" & quatrepoints$grille=="100100" &
quatrepoints$population=="validdiff" & quatrepoints$pop=="cowTest" )
```

```
summary(construction)
```

```

##sélection du meilleur modèle cherchant à expliquer les variations observées de la NEC en fonction des
## coordonnées des surfaces 3D dans la population de calibration dans le repère de l'ACP: ##

adjrcarre<-
data.frame(modele=0,étape=0,rcarré=0,rcarrévalididem=0,rcarrévaliddiff=0,rcarrépoints=0,rcarréimage=0,erreur=0,erreurva
lididem=0,erreurvaliddiff=0,erreurpoints=0,erreurimage=0)  ###création d'une table contenant #
# pour chaque étape du stepwise le r2 ajusté du modèle et le r2 ajusté du modèle sur la population validation externe
##(rcarrévalididem et rcarrévaliddiff), sur les populations pour la répétabilité (rcarrépoints et
## rcarréimage) et l'étape du stepwise. Cette étape renseigne que pour démarrer le modèle ne contient
## aucune variable, adjr2=0

for (i in 1:57)
{stepi<-step(lm(nec~1,construction),nec ~
cp1+cp2+cp3+cp4+cp5+cp6+cp7+cp8+cp9+cp10+cp11+cp12+cp13+cp14+cp15+cp16+cp17+cp18+cp19+cp20+cp21+cp22+cp23+cp24+cp25+cp2
6+cp27+cp28+cp29+cp30+cp31+cp32+cp33+cp34+cp35+cp36+cp37+cp38+cp39+cp40+cp41+cp42+cp43+cp44+cp45+cp46+cp47+cp48+cp49+cp
50+cp51+cp52+cp53+cp54+cp55+cp56+cp57,steps=i,direction="both",data=construction)
outputi<-summary(stepi)
adjrsqi<-outputi$adj.r.squared
erreur1<-sd(outputi$residuals)
## population de validation avec les individus identiques à la population de construction: ##
prediction_idem<-predict(stepi,nec3d_idem,se.fit=TRUE)
predict3didem<-as.data.frame(prediction_idem$fit)
stepvalid_idem<-cbind(nec3d_idem,predict3didem)
valididem.stepi<-lm(nec~prediction_idem$fit,data=stepvalid_idem)
adjrsqvalididem<-summary(valididem.stepi)$adj.r.squared
erreurvalidid<-sd(summary(valididem.stepi)$residuals)
## population de validation avec les individus différents de la population de construction: ##
prediction_diff<-predict(stepi,nec3d_diff,se.fit=TRUE)
predict3ddiff<-as.data.frame(prediction_diff$fit)
stepvalid_diff<-cbind(nec3d_diff,predict3ddiff)
validdiff.stepi<-lm(nec~prediction_diff$fit,data=stepvalid_diff)
adjrsqvaliddiff<-summary(validdiff.stepi)$adj.r.squared
erreurvaliddif<-sd(summary(validdiff.stepi)$residuals)
## population pour tester la répétabilité du pointage: ##
prediction_pts<-predict(stepi,pointage,se.fit=TRUE)
predict3dpts<-as.data.frame(prediction_pts$fit)
steppts<-cbind(pointage,predict3dpts)
pts.stepi<-lm(nec~prediction_pts$fit,data=steppts)
adjrsqpts<-summary(pts.stepi)$adj.r.squared
erreurpts<-sd(summary(pts.stepi)$residuals)
## population pour tester la répétabilité de l'indicateur sur un même individu pour des images
## différentes:

```

```

prediction_im<-predict(stepi,image,se.fit=TRUE)
predict3dim<-as.data.frame(prediction_im$fit)
stepim<-cbind(image,predict3dim)
im.stepi<-lm(nec~prediction_im$fit,data=stepim)
adjrsqim<-summary(im.stepi)$adj.r.squared
erreurim<-sd(summary(im.stepi)$residuals)
## construction d'une base de données contenant tous les R2 de tous les modèles pour chacune des
## populations:
adjrcarrei<-
data.frame(modele=0,étape=i,rcarré=adjrsqi,rcarrévalididem=adjrsqvalididem,rcarrévaliddiff=adjrsqvaliddiff,rcarrépoints
=adjrsqpts,rcarréimage=adjrsqim,erreur=erreur1,erreurvalididem=erreurvalidid,erreurvaliddiff=erreurvaliddif,erreurpoint
s=erreurpts,erreurimage=erreurim)
adjrcarre<-rbind(adjrcarre,adjrcarrei) ## fusionne l'ancienne table r2 avec la nouvelle ligne calculée
## pour l'étape i
}

### représentation graphique de l'évolution des modèles en fonction du nombre de composantes dans le modèle, r2=f(i):
##
modelebest<-stepi$call ## extrait le modèle complet du step
modelebest
stepi
modell<-gsub("nec ~ ","",modelebest, fixed=TRUE)
## remplacement exactement les caractères "nec ~ " du modèle par un vide "" avec l'option fixed=TRUE
## qui permet de remplacer exactement les caractères
modell
model<-gsub("+","",modell[2],fixed=TRUE)
model
composante<-strsplit(model," ") ## coupe la chaîne de caractères au niveau des " "
composante<-noquote(composante[[1]])## retire les trêmas de la chaîne de caractères
composante
composante<-as.matrix(as.factor(composante))
composante
dim(composante)
cp<-as.data.frame((as.matrix(composante))) ## transformation en data.frame
dim(cp) ## donne le nombre de cp du meilleur modèle
adjrcarre
listemodele<-cbind(listemodele,cp)

write.table(adjrcarre,file=paste("F:/resultats_nec3d/calibration/",paste(construction$methode[1],construction$grille[1]
),sep=""),sep=";",dec=".",append=FALSE,quote=TRUE)

```

```

### traçons les graphes pour voir l'évolution des modèles au fûr et à mesure de l'avancement du stepwise: windows() ####
ouvre la console affichant les graphes sous R ## barplot(adjrcarre$rcarré,names.arg=adjrcarre$étape,main="évolution de
la qualité des modèles du stepwise",xlab="nombre d'étapes dans la sélection par stepwise",ylab="R-carré
ajusté",ylim=c(0,1))
savePlot(filename=paste("F:/resultats_nec3d/calibration/",paste(construction$methode[1],construction$grille[1],construc
tion$population[1],construction$pop[1]),sep=""),type="jpeg",device=dev.cur())

```

```

### enregistre le graphe sous format .jpeg sur mon disque dur externe (F:). La commande path.expand
### renseigne le chemin d'accès où R doit enregistrer le document!
barplot(adjrcarre$rcarrévalididem,names.arg=adjrcarre$étape,main=paste("évolution de la qualité des modèles du
stepwise","appliqués sur la population de validation","avec les vaches identiques à la population de
construction",sep="\n"),xlab="nombre d'étapes dans la sélection par stepwise",ylab="R-carré ajusté",ylim=c(0,1))
savePlot(filename=paste("F:/resultats_nec3d/calibration/",paste(nec3d_idem$methode[1],nec3d_idem$grille[1],nec3d_idem$popu
lation[1]),sep=""),type="jpeg",device=dev.cur())
barplot(adjrcarre$rcarrévaliddiff,names.arg=adjrcarre$étape,main=paste("évolution de la qualité des modèles du
stepwise","appliqués sur la population de validation", "avec les vaches différentes de la population de
construction",sep="\n"), xlab="nombre d'étapes dans la sélection par stepwise",ylab="R-carré ajusté",ylim=c(0,1))
savePlot(filename=paste("F:/resultats_nec3d/calibration/",paste(nec3d_diff$methode[1],nec3d_diff$grille[1],nec3d_diff$popu
lation[1],nec3d_diff$pop[1]),sep=""),type="jpeg",device=dev.cur())
barplot(adjrcarre$rcarrépoints,names.arg=adjrcarre$étape,main=paste("évolution de la qualité des modèles du
stepwise","appliqués sur la population test pour la répétabilité du pointage",sep="\n"),xlab="nombre d'étapes dans la
sélection par stepwise",ylab="R-carré ajusté",ylim=c(0,1))
savePlot(filename=paste("F:/resultats_nec3d/calibration/",paste(pointage$methode[1],pointage$grille[1],pointage$populat
ion[1]),sep=""),type="jpeg",device=dev.cur())
barplot(adjrcarre$rcarréimage,names.arg=adjrcarre$étape,main=paste("évolution de la qualité des modèles du
stepwise","appliqués sur la population test pour la répétabilité intra-individu",sep="\n"),xlab="nombre d'étapes dans
la sélection par stepwise",ylab="R-carré ajusté",ylim=c(0,1))
savePlot(filename=paste("F:/resultats_nec3d/calibration/",paste(image$methode[1],image$grille[1],image$population[1]),s
ep=""),type="jpeg",device=dev.cur())

```

```

## graphe pour les erreurs de prédiction:

```

```

barplot(adjrcarre$erreur,names.arg=adjrcarre$étape,main="évolution de la qualité des modèles du stepwise",xlab="nombre
d'étapes dans la sélection par stepwise",ylab="erreur de prediction",ylim=c(0,0.8))
savePlot(filename=paste("F:/resultats_nec3d/calibration/", "erreur",paste(colnames(adjrcarre[,8]),construction$methode[1
],construction$grille[1],construction$population[1],construction$pop[1]),sep=""),type="jpeg",device=dev.cur())

```

```

### enregistre le graphe sous format .jpeg sur mon disque dur externe (F:). La commande path.expand

```

```

## renseigne le chemin d'accès où R doit enregistrer le document!

```

```

barplot(adjrcarre$erreurvalididem,names.arg=adjrcarre$étape,main=paste("évolution de la qualité des modèles du
stepwise","appliqués sur la population de validation","avec les vaches identiques à la population de
construction",sep="\n"),xlab="nombre d'étapes dans la sélection par stepwise",ylab="erreur de
prediction",ylim=c(0,0.8))

```

```

savePlot(filename=paste("F:/resultats_nec3d/calibration/", "erreur", paste(colnames(adjrcarre[,9]), nec3d_idem$methode[1],
nec3d_idem$grille[1], nec3d_idem$population[1]), sep=" "), type="jpeg", device=dev.cur())
barplot(adjrcarre$erreurvaliddiff, names.arg=adjrcarre$etape, main=paste("évolution de la qualité des modèles du
stepwise", "appliqués sur la population de validation", "avec les vaches différentes de la population de
construction", sep="\n"), xlab="nombre d'étapes dans la sélection par stepwise", ylab="erreur de
prediction", ylim=c(0,0.8))
savePlot(filename=paste("F:/resultats_nec3d/calibration/", "erreur", paste(colnames(adjrcarre[,10]), nec3d_diff$methode[1]
, nec3d_diff$grille[1], nec3d_diff$population[1], nec3d_diff$pop[1]), sep=" "), type="jpeg", device=dev.cur())
barplot(adjrcarre$erreurpoints, names.arg=adjrcarre$etape, main=paste("évolution de la qualité des modèles du
stepwise", "appliqués sur la population test pour la répétabilité du pointage", sep="\n"), xlab="nombre d'étapes dans la
sélection par stepwise", ylab="erreur de prediction", ylim=c(0,0.8))
savePlot(filename=paste("F:/resultats_nec3d/calibration/", "erreur", paste(colnames(adjrcarre[,11]), pointage$methode[1], p
ointage$grille[1], pointage$population[1]), sep=" "), type="jpeg", device=dev.cur())
barplot(adjrcarre$erreurimage, names.arg=adjrcarre$etape, main=paste("évolution de la qualité des modèles du
stepwise", "appliqués sur la population test pour la répétabilité intra-individu", sep="\n"), xlab="nombre d'étapes dans
la sélection par stepwise", ylab="erreur de prediction", ylim=c(0,0.8))
savePlot(filename=paste("F:/resultats_nec3d/calibration/", "erreur", paste(colnames(adjrcarre[,12]), image$methode[1], imag
e$grille[1], image$population[1]), sep=" "), type="jpeg", device=dev.cur())

```

```

windows()
plot(stepi$fitted.values, construction$nec, type="n", main=paste("qualité du modèle de
calibration", paste(construction$methode[1], construction$grille[1], sep=" "), sep=" "), xlab="nec prédite", ylab="nec
observées", xlim=c(0,5), ylim=c(0,5))
abline(lm(construction$nec~stepi$fitted.values))
abline(a=0, b=1, col="red", lty=2)
savePlot(filename=paste("F:/resultats_nec3d/calibration/", "modelecalib", paste(construction$methode[1], construction$gril
le[1], sep=" "), sep=" "), type="jpeg", device=dev.cur())

```

```

#####
##### processus de caractérisation du modèle: répétabilité et reproductibilité #####

```

```
##vérifions que les tableaux de données sont conformes: ##
```

```

summary(pointage)
rownames(pointage)<-pointage[,1]
rownames(image)<-image[,1]
summary(image)

```

```

## testons la répétabilité de l'identification des points: ##
## pour cela, il faut prédire la nec3d à partir du modèle retenu à l'issue de la stepwise, modelebest

```

```

prediction_pts<-predict(stepi,pointage,se.fit=TRUE)
## estimation des nec3d de la population à partir du modèle conservé ##
valeurspredites_pts<-prediction_pts$fit
valeurspredites_pts
## les sorties sont sous forme d'un vecteur, transformons le en tableau et fusionnons le avec la base
## de données pointage:
nec3dpts<-as.data.frame(valeurspredites_pts)
test_pointage<-cbind(pointage,nec3dpts)
test_pointage

## calculons la répétabilité de l'indicateur: ##
fidelite<-matrix("",nrow=2,ncol=6)
fidelite[1,1]<-"répétabilité"
fidelite[2,1]<-"reproductibilité"

## les identifiants des vaches utilisés se confondent avec les séries, on supprime le premier chiffre:
id<-substring(pointage$vache,2,4)
id
id<-gsub("575","075",id,fixed=TRUE)
id<-as.factor(id)
id
test_pointage<-cbind(test_pointage,id)
test_pointage

## réalisation de l'anova pour analyser l'erreur liée au pointage, c'est-à-dire en retirant l'effet de l'identifiant
## de la vache:
anova_pts<-aov(valeurspredites_pts~id,data=test_pointage)
summary(anova_pts)
dllr=48-anova_pts$rank+1
sr2=var(anova_pts$residuals)*47/dllr
sr2
M=mean(valeurspredites_pts)
r=sqrt(sr2)
r
## pour les mêmes conditions expérimentales,même ferme, même personne, même méthode, la différence
## maximale au risque de ## 5% qui sépare 2 résultats sur une même acquisition vaut ##
r
## unité de NEC sur un échantillon dont la moyenne est proche de ##
M #.#
cvr=r/M*100 ## coefficient de variation de répétabilité ##
cvr

fidelite[1,2]<-dllr

```

```

fidelite[1,3]<-sr2
fidelite[1,4]<-r
fidelite[1,5]<-M
fidelite[1,6]<-cvr

## testons la reproductibilité intra-ferme de l'indicateur: ##

prediction_im<-predict(stepi,image,se.fit=TRUE)  ## estimation des nec3d de la population à partir du
                                                ## modèle conservé ##

valeurspredites_im<-prediction_im$fit
valeurspredites_im
## les sorties sont sous forme d'un vecteur, transformons le en tableau et fusionnons le avec la base de
## données pointage:
nec3dim<-as.data.frame(valeurspredites_im)
test_image<-cbind(image,nec3dim)
test_image

## calculons la répétabilité intra-ferme de l'indicateur: ##
idr<-substring(image$vache,2,4)
idr
idr<-gsub("575","075",idr,fixed=TRUE)
idr<-as.factor(idr)
idr
test_image<-cbind(test_image,idr)
test_image

anova_im<-aov(valeurspredites_im~idr,data=test_image)
summary(anova_im) dllrim=48-
anova_im$rank+1
srim2=var(anova_im$residuals)*47/dllrim
srim2
Mim=mean(valeurspredites_im)
rim=sqrt(srim2)
rim

## pour les mêmes conditions expérimentales,même ferme, même personne, même méthode, la différence
## maximale au risque de 5% qui sépare 2 résultats sur une même acquisition vaut ##
rim
# unité de NEC sur un échantillon dont la moyenne est proche de #
Mim #.#
cvrim=rim/Mim*100 ##coefficient de variation de répétabilité ##
cvrim

```



```

fidelite[2,2]<-dllrim
fidelite[2,3]<-srim2
fidelite[2,4]<-rim
fidelite[2,5]<-Mim
fidelite[2,6]<-cvrim

fidelite<-as.data.frame(fidelite)
colnames(fidelite)<-c("fidélité","dllr","srcarré","répétabilité ou reproductibilité","moyenne","CV")
write.table(fidelite,file=paste("F:/resultats_nec3d/fidelite/",paste(pointage$methode[1],pointage$grille[1]),sep=""),sep=";",dec="," ,row.names=FALSE)

```

```

#####
#####
##### processus de validation #####

```

```
## vérifions les bases de données des populations de validation: ##
```

```

summary(nec3d_idem)
rownames(nec3d_idem)<-nec3d_idem[,1]
summary(nec3d_diff)
rownames(nec3d_diff)<-nec3d_diff[,1]

```

```
## validation sur la population de validation contenant les mêmes vaches que la populaion de construction:
```

```

prediction_idem<-predict(stepi,nec3d_idem,se.fit=TRUE)
nec3didem<-prediction_idem$fit
nec3didem
nec3did<-as.data.frame(nec3didem)
valid_idem<-cbind(nec3d_idem,nec3did)
valid_idem

```

```
## qualité de l'estimation pour des vaches identiques à la population de construction mais avec des
## acquisitions différentes: ##
```

```

modele_valid_idem<-lm(nec~nec3didem,data=valid_idem)
rcarre_valididem<-summary(modele_valid_idem)$adj.r.squared
varianceerreur_valididem<-summary(modele_valid_idem)$sigma

```

```

validdiff<-matrix("",nrow=2,ncol=7)
validdiff[1,1]<-"valididem"
validdiff[2,1]<-"validdiff"

## calcul des différents critères pour la validation:
biais2_3id<-(mean(valid_idem$nec)-mean(valid_idem$nec3didem))^2
NU3id<-(1-modele_valid_idem$coefficients[2])^2*var(valid_idem$nec3didem)
LC3id<-(1-cor(valid_idem$nec,valid_idem$nec3didem)^2)*var(valid_idem$nec)
MSE3id<-biais2_3id+NU3id+LC3id
effetbiais_3id<-biais2_3id/MSE3id
effetpenete_3id<-NU3id/MSE3id
effetresidu_3id<-LC3id/MSE3id
validdiff[1,2]<-MSE3id
validdiff[1,3]<-effetbiais_3id
validdiff[1,4]<-effetpenete_3id
validdiff[1,5]<-effetresidu_3id
validdiff[1,6]<-summary(modele_valid_idem)$r.squared
validdiff[1,7]<-sd(summary(modele_valid_idem)$residuals)

##### validation sur la population de validation contenant des vaches #####
##### différents de la population de construction: #####

prediction_diff<-predict(stepi,nec3d_diff,se.fit=TRUE)
nec3ddiff<-prediction_diff$fit
nec3ddiff
nec3ddi<-as.data.frame(nec3ddiff)
valid_diff<-cbind(nec3d_diff,nec3ddi)
valid_diff

## qualité de l'estimation pour des vaches autres que celles de la population de construction mais avec
## des acquisitions différentes: ##

modele_valid_diff<-lm(nec~nec3ddiff,data=valid_diff)
rcarre_validdiff<-summary(modele_valid_diff)$adj.r.squared
rcarre_validdiff
varianceerreur_validdiff<-summary(modele_valid_diff)$sigma
varianceerreur_validdiff

biais2_3dif<-(mean(valid_diff$nec)-mean(valid_diff$nec3ddiff))^2
NU3dif<-(1-modele_valid_diff$coefficients[2])^2*var(valid_diff$nec3ddiff)
LC3dif<-(1-cor(valid_diff$nec,valid_diff$nec3ddiff)^2)*var(valid_diff$nec)
MSE3dif<-biais2_3dif+NU3dif+LC3dif

```

```

effetbiais_3dif<-biais2_3dif/MSE3dif
effetpenste_3dif<-NU3dif/MSE3dif
effetresidu_3dif<-LC3dif/MSE3dif
validdiff[2,2]<-MSE3dif
validdiff[2,3]<-effetbiais_3dif
validdiff[2,4]<-effetpenste_3dif
validdiff[2,5]<-effetresidu_3dif
validdiff[2,6]<-summary(modele_valid_diff)$r.squared
validdiff[2,7]<-sd(summary(modele_valid_diff)$residuals)

validdiff<-as.data.frame(validdiff)
colnames(validdiff)<-c("validation","biaisglobal","biaismoyenne","biaispenste","biaisinexpliqué","rcarré","erreur
prediction")
fichiervalidation<-paste(validdiff$methode[1],validdiff$grille[1])
write.table(validdiff,file=paste("F:/resultats_nec3d/validation/",paste(valid_diff$methode[1],valid_diff$grille[1],sep=
""),sep=""),sep=";",dec=" ",row.names=FALSE)

windows()
plot(valid_diff$nec3ddiff,valid_diff$nec,type="p",main="validation de nec3d sur la population
validdiff",col=4,xlim=c(0,5),ylim=c(0,5))
abline(modele_valid_diff,col=4)
abline(a=0,b=1,lty=2,col=1)
savePlot(filename=paste("F:/resultats_nec3d/validation/",paste(valid_diff$methode[1],valid_diff$grille[1],valid_diff$po
pulation[1],sep=""),sep=""),type="jpeg",device=dev.cur())

plot(valid_idem$nec3didem,valid_idem$nec,type="p",main="validation de nec3d sur la population
valididem",col=2,xlim=c(0,5),ylim=c(0,5))
abline(modele_valid_idem,col=2) ## représentation graphique du modèle abline(a=0,b=1,lty=2,col=1) ## représentation
graphique de la première bissectrice
savePlot(filename=paste("F:/resultats_nec3d/validation/",paste(valid_idem$methode[1],valid_idem$grille[1],valid_idem$po
pulation[1],sep=""),sep=""),type="jpeg",device=dev.cur())

#####
#####
##### extraction des résidus des modèles appliqués sur les 2 populations de validation #####

## extraction des résidus pour chacune des populations de validation:

residus_validdiff<-as.data.frame(modele_valid_diff$residuals)

```

```

residus_validdiff<-as.data.frame(residus_validdiff[,1])
colnames(residus_validdiff)<-paste(valid_diff$methode[1],valid_diff$grille[1],sep="")

residus_valididem<-as.data.frame(modele_valid_idem$residuals)
residus_valididem<-as.data.frame(residus_valididem[,1])
colnames(residus_valididem)<-paste(valid_idem$methode[1],valid_idem$grille[1],sep="")

resdiff<-cbind(resdiff,residus_validdiff)
write.table(resdiff,file=paste("F:/resultats_nec3d/comparaison/", "residusdiff.csv", sep=""), sep=";", dec=".", row.names=FA
LSE)
residem<-cbind(residem,residus_valididem)
write.table(residem,file=paste("F:/resultats_nec3d/comparaison/", "residusidem.csv", sep=""), sep=";", dec=".", row.names=FA
LSE)

```

```

#####
#####
#####
#####
##### méthode 4 points imitant 3 points pour une grille 50x50 #####
#####

```

```

construction<-subset( quatrepoints, quatrepoints$methode=="4points3remplissage" & quatrepoints$grille==" 5050" &
quatrepoints$pop=="cowDB" )
pointage<-subset( quatrepoints, quatrepoints$methode=="4points3remplissage" & quatrepoints$grille==" 5050" &
quatrepoints$population=="pointage" )
image<-subset( quatrepoints, quatrepoints$methode=="4points3remplissage" & quatrepoints$grille==" 5050" &
quatrepoints$population=="image" )
nec3d_idem<-subset( quatrepoints, quatrepoints$methode=="4points3remplissage" & quatrepoints$grille==" 5050" &
quatrepoints$population=="valididem" )
nec3d_diff<-subset( quatrepoints, quatrepoints$methode=="4points3remplissage" & quatrepoints$grille==" 5050" &
quatrepoints$population=="validdiff" & quatrepoints$pop=="cowTest" )

```

```
summary(construction)
```

```
##sélection du meilleur modèle cherchant à expliquer les variations observées de la NEC en fonction des
## coordonnées des surfaces 3D dans la population de calibration dans le repère de l'ACP: ##
```

```

adjrcarre<-
data.frame(modele=0,étape=0,rcarré=0,rcarrévalididem=0,rcarrévaliddiff=0,rcarrépoints=0,rcarréimage=0,erreur=0,erreurva
lididem=0,erreurvaliddiff=0,erreurpoints=0,erreurimage=0) ###création d'une table contenant #

```

```

# pour chaque étape du stepwise le r2 ajusté du modèle et le r2 ajusté du modèle sur la population validation
## externe (rcarrévalididem et rcarrévaliddiff), sur les populations pour la répétabilité (rcarrépoints et
## rcarréimage) et l'étape du stepwise. Cette étape renseigne que pour démarrer le modèle ne contient aucune
## variable, adjr2=0 ##

for (i in 1:57)
{stepi<-step(lm(nec~1,construction),nec ~
cp1+cp2+cp3+cp4+cp5+cp6+cp7+cp8+cp9+cp10+cp11+cp12+cp13+cp14+cp15+cp16+cp17+cp18+cp19+cp20+cp21+cp22+cp23+cp24+cp25+cp2
6+cp27+cp28+cp29+cp30+cp31+cp32+cp33+cp34+cp35+cp36+cp37+cp38+cp39+cp40+cp41+cp42+cp43+cp44+cp45+cp46+cp47+cp48+cp49+cp
50+cp51+cp52+cp53+cp54+cp55+cp56+cp57,steps=i,direction="both",data=construction)
outputi<-summary(stepi)
adjrsqi<-outputi$adj.r.squared
erreur1<-sd(outputi$residuals)
## population de validation avec les individus identiques à la population de construction: ##
prediction_idem<-predict(stepi,nec3d_idem,se.fit=TRUE)
predict3didem<-as.data.frame(prediction_idem$fit)
stepvalid_idem<-cbind(nec3d_idem,predict3didem)
valididem.stepi<-lm(nec~prediction_idem$fit,data=stepvalid_idem)
adjrsqvalididem<-summary(valididem.stepi)$adj.r.squared
erreurvalidid<-sd(summary(valididem.stepi)$residuals)
## population de validation avec les individus différents de la population de construction: ##
prediction_diff<-predict(stepi,nec3d_diff,se.fit=TRUE)
predict3ddiff<-as.data.frame(prediction_diff$fit)
stepvalid_diff<-cbind(nec3d_diff,predict3ddiff)
validdiff.stepi<-lm(nec~prediction_diff$fit,data=stepvalid_diff)
adjrsqvaliddiff<-summary(validdiff.stepi)$adj.r.squared
erreurvaliddif<-sd(summary(validdiff.stepi)$residuals)
## population pour tester la répétabilité du pointage: ##
prediction_pts<-predict(stepi,pointage,se.fit=TRUE)
predict3dpts<-as.data.frame(prediction_pts$fit)
steppts<-cbind(pointage,predict3dpts)
pts.stepi<-lm(nec~prediction_pts$fit,data=steppts)
adjrsqpts<-summary(pts.stepi)$adj.r.squared
erreurpts<-sd(summary(pts.stepi)$residuals)
## population pour tester la répétabilité de l'indicateur sur un même individu pour des images différentes:
prediction_im<-predict(stepi,image,se.fit=TRUE)
predict3dim<-as.data.frame(prediction_im$fit)
stepim<-cbind(image,predict3dim)
im.stepi<-lm(nec~prediction_im$fit,data=stepim)
adjrsqim<-summary(im.stepi)$adj.r.squared
erreurim<-sd(summary(im.stepi)$residuals)
## construction d'une base de données contenant tous les R2 de tous les modèles pour chacune

```

```

## des populations:
adjrcarrei<-
data.frame(modele=0,étape=i,rcarré=adjrsqi,rcarrévalididem=adjrsqvalididem,rcarrévaliddiff=adjrsqvaliddiff,rcarrépoints
=adjrsqpts,rcarréimage=adjrsqim,erreur=erreur1,erreurvalididem=erreurvalidid,erreurvaliddiff=erreurvaliddif,erreurpoint
s=erreurpts,erreurimage=erreurim)
adjrcarre<-rbind(adjrcarre,adjrcarrei) ## fusionne l'ancienne table r2 avec la nouvelle ligne calculée
## pour l'étape i
}

### représentation graphique de l'évolution des modèles en fonction du nombre de composantes dans le
### modèle, r2=f(i):
modelebest<-stepi$call ## extrait le modèle complet du step
modelebest
stepi
modell<-gsub("nec ~ ","",modelebest, fixed=TRUE)
## remplacement exactement les caractères "nec ~ " du modèle par un vide "" avec l'option fixed=TRUE
## qui permet de remplacer exactement les caractères
modell
model<-gsub("+","",modell[2],fixed=TRUE)
model
composante<-strsplit(model," ") ## coupe la chaîne de caractères au niveau des " "
composante<-noquote(composante[[1]])## retire les trêmas de la chaîne de caractères
composante
composante<-as.matrix(as.factor(composante))
composante
dim(composante)
cp<-as.data.frame((as.matrix(composante))) ## transformation en data.frame
dim(cp) ## donne le nombre de cp du meilleur modèle
adjrcarre
listemodele<-cbind(listemodele,cp)

write.table(adjrcarre,file=paste("F:/resultats_nec3d/calibration/",paste(construction$methode[1],construction$grille[1]
),sep=""),sep=";",dec=" ",append=FALSE,quote=TRUE)

### traçons les graphes pour voir l'évolution des modèles au fûr et à mesure de l'avancement du stepwise:
windows()
barplot(adjrcarre$rcarré,names.arg=adjrcarre$étape,main="évolution de la qualité des modèles du stepwise",xlab="nombre
d'étapes dans la sélection par stepwise",ylab="R-carré ajusté",ylim=c(0,1))
savePlot(filename=paste("F:/resultats_nec3d/calibration/",paste(construction$methode[1],construction$grille[1],construc
tion$population[1],construction$pop[1]),sep=""),type="jpeg",device=dev.cur())
### enregistre le graphe sous format .jpeg sur mon disque dur externe (F:). La commande path.expand
## renseigne le chemin d'accès où R doit enregistrer le document! ###

```

```

barplot(adjrcarre$rcarrévalididem,names.arg=adjrcarre$étape,main=paste("évolution de la qualité des modèles du
stepwise","appliqués sur la population de validation","avec les vaches identiques à la population de
construction",sep="\n"),xlab="nombre d'étapes dans la sélection par stepwise",ylab="R-carré ajusté",ylim=c(0,1))
savePlot(filename=paste("F:/resultats_nec3d/calibration/",paste(nec3d_idem$methode[1],nec3d_idem$grille[1],nec3d_idem$popu
lation[1]),sep=""),type="jpeg",device=dev.cur())
barplot(adjrcarre$rcarrévaliddiff,names.arg=adjrcarre$étape,main=paste("évolution de la qualité des modèles du
stepwise","appliqués sur la population de validation", "avec les vaches différentes de la population de
construction",sep="\n"), xlab="nombre d'étapes dans la sélection par stepwise",ylab="R-carré ajusté",ylim=c(0,1))
savePlot(filename=paste("F:/resultats_nec3d/calibration/",paste(nec3d_diff$methode[1],nec3d_diff$grille[1],nec3d_diff$popu
lation[1],nec3d_diff$pop[1]),sep=""),type="jpeg",device=dev.cur())
barplot(adjrcarre$rcarrépoints,names.arg=adjrcarre$étape,main=paste("évolution de la qualité des modèles du
stepwise","appliqués sur la population test pour la répétabilité du pointage",sep="\n"),xlab="nombre d'étapes dans la
sélection par stepwise",ylab="R-carré ajusté",ylim=c(0,1))
savePlot(filename=paste("F:/resultats_nec3d/calibration/",paste(pointage$methode[1],pointage$grille[1],pointage$populat
ion[1]),sep=""),type="jpeg",device=dev.cur())
barplot(adjrcarre$rcarréimage,names.arg=adjrcarre$étape,main=paste("évolution de la qualité des modèles du
stepwise","appliqués sur la population test pour la répétabilité intra-individu",sep="\n"),xlab="nombre d'étapes dans
la sélection par stepwise",ylab="R-carré ajusté",ylim=c(0,1))
savePlot(filename=paste("F:/resultats_nec3d/calibration/",paste(image$methode[1],image$grille[1],image$population[1]),s
ep=""),type="jpeg",device=dev.cur())

```

graphe pour les erreurs de prédiction:

```

barplot(adjrcarre$erreur,names.arg=adjrcarre$étape,main="évolution de la qualité des modèles du stepwise",xlab="nombre
d'étapes dans la sélection par stepwise",ylab="erreur de prediction",ylim=c(0,0.8))
savePlot(filename=paste("F:/resultats_nec3d/calibration/","erreur",paste(colnames(adjrcarre[,8]),construction$methode[1]
,construction$grille[1],construction$population[1],construction$pop[1]),sep=""),type="jpeg",device=dev.cur())

```

enregistre le graphe sous format .jpeg sur mon disque dur externe (F:). La commande path.expand

renseigne le chemin d'accès où R doit enregistrer le document!

```

barplot(adjrcarre$erreurvalididem,names.arg=adjrcarre$étape,main=paste("évolution de la qualité des modèles du
stepwise","appliqués sur la population de validation","avec les vaches identiques à la population de
construction",sep="\n"),xlab="nombre d'étapes dans la sélection par stepwise",ylab="erreur de
prediction",ylim=c(0,0.8))
savePlot(filename=paste("F:/resultats_nec3d/calibration/","erreur",paste(colnames(adjrcarre[,9]),nec3d_idem$methode[1],
nec3d_idem$grille[1],nec3d_idem$population[1]),sep=""),type="jpeg",device=dev.cur())
barplot(adjrcarre$erreurvaliddiff,names.arg=adjrcarre$étape,main=paste("évolution de la qualité des modèles du
stepwise","appliqués sur la population de validation", "avec les vaches différentes de la population de
construction",sep="\n"), xlab="nombre d'étapes dans la sélection par stepwise",ylab="erreur de
prediction",ylim=c(0,0.8))
savePlot(filename=paste("F:/resultats_nec3d/calibration/","erreur",paste(colnames(adjrcarre[,10]),nec3d_diff$methode[1]
,nec3d_diff$grille[1],nec3d_diff$population[1],nec3d_diff$pop[1]),sep=""),type="jpeg",device=dev.cur())

```

```

barplot(adjrcarre$erreurpoints,names.arg=adjrcarre$étape,main=paste("évolution de la qualité des modèles du
stepwise","appliqués sur la population test pour la répétabilité du pointage",sep="\n"),xlab="nombre d'étapes dans la
sélection par stepwise",ylab="erreur de prediction",ylim=c(0,0.8))
savePlot(filename=paste("F:/resultats_nec3d/calibration/","erreur",paste(colnames(adjrcarre[,11]),pointage$methode[1],p
ointage$grille[1],pointage$population[1]),sep=""),type="jpeg",device=dev.cur())
barplot(adjrcarre$erreurimage,names.arg=adjrcarre$étape,main=paste("évolution de la qualité des modèles du
stepwise","appliqués sur la population test pour la répétabilité intra-individu",sep="\n"),xlab="nombre d'étapes dans
la sélection par stepwise",ylab="erreur de prediction",ylim=c(0,0.8))
savePlot(filename=paste("F:/resultats_nec3d/calibration/","erreur",paste(colnames(adjrcarre[,12]),image$methode[1],imag
e$grille[1],image$population[1]),sep=""),type="jpeg",device=dev.cur())

```

```

windows()
plot(stepi$fitted.values,construction$nec,type="n",main=paste("qualité du modèle de
calibration",paste(construction$methode[1],construction$grille[1],sep=""),sep=""),xlab="nec prédite",ylab="nec
observées",xlim=c(0,5),ylim=c(0,5))
abline(lm(construction$nec~stepi$fitted.values)) abline(a=0,b=1,col="red",lty=2)
savePlot(filename=paste("F:/resultats_nec3d/calibration/","modelecalib",paste(construction$methode[1],construction$gril
le[1],sep=""),sep=""),type="jpeg",device=dev.cur())

```

```

#####
##### processus de caractérisation du modèle: répétabilité et reproductibilité #####
#####

```

```
##vérifions que les tableaux de données sont conformes: ##
```

```

summary(pointage)
rownames(pointage)<-pointage[,1]
rownames(image)<-image[,1]
summary(image)

```

```

## testons la répétabilité de l'identification des points: ##
## pour cela, il faut prédire la nec3d à partir du modèle retenu à l'issu de la stepwise, modelebest
## prediction_pts<-predict(stepi,pointage,se.fit=TRUE) ## estimation des nec3d de la population à partir du modèle
## conservé ##
valeurspredites_pts<-prediction_pts$fit
valeurspredites_pts
## les sorties sont sous forme d'un vecteur, transformons le en tableau et fusionnons le avec la base de
## données «pointage»:
nec3dpts<-as.data.frame(valeurspredites_pts)
test_pointage<-cbind(pointage,nec3dpts)

```



```

test_pointage

## calculons la répétabilité de l'indicateur: ##
fidelite<-matrix("",nrow=2,ncol=6)
fidelite[1,1]<-"répétabilité"
fidelite[2,1]<-"reproductibilité"

## les identifiants des vaches utilisés se confondent avec les séries, on supprime le premier chiffre:
id<-substring(pointage$vache,2,4)
id
id<-gsub("575","075",id,fixed=TRUE)
id<-as.factor(id)
id
test_pointage<-cbind(test_pointage,id)
test_pointage

## réalisation de l'anova pour analyser l'erreur liée au pointage, c'est-à-dire en retirant l'effet de
## l'identifiant de la vache:
anova_pts<-aov(valeurspredites_pts~id,data=test_pointage)
summary(anova_pts) dllr=48-
anova_pts$rank+1
sr2=var(anova_pts$residuals)*47/dllr
sr2
M=mean(valeurspredites_pts)
r=sqrt(sr2)
r
## pour les mêmes conditions expérimentales,même ferme, même personne, même méthode, la différence maximale
## au risque de 5% qui sépare 2 résultats sur une même acquisition vaut ##
r
## unité de NEC sur un échantillon dont la moyenne est proche de ##
M #.#
cvr=r/M*100 ## coefficient de variation de répétabilité ##
cvr

fidelite[1,2]<-dllr
fidelite[1,3]<-sr2
fidelite[1,4]<-r
fidelite[1,5]<-M
fidelite[1,6]<-cvr

## testons la reproductibilité intra-ferme de l'indicateur: ##

```

```

prediction_im<-predict(stepi,image,se.fit=TRUE) ## estimation des nec3d de la population à partir du modèle conservé
valeurspredites_im<-prediction_im$fit
valeurspredites_im
## les sorties sont sous forme d'un vecteur, transformons le en tableau et fusionnons le avec la base de
## données pointage:
nec3dim<-as.data.frame(valeurspredites_im)
test_image<-cbind(image,nec3dim)
test_image

## calculons la répétabilité intra-ferme de l'indicateur: ##
idr<-substring(image$vache,2,4)
idr
idr<-gsub("575","075",idr,fixed=TRUE)
idr<-as.factor(idr)
idr
test_image<-cbind(test_image,idr)
test_image
anova_im<-aov(valeurspredites_im~idr,data=test_image)
summary(anova_im) dllrim=48-
anova_im$rank+1
srim2=var(anova_im$residuals)*47/dllrim
srim2
Mim=mean(valeurspredites_im)
rim=sqrt(srim2)
rim

## pour les mêmes conditions expérimentales,même ferme, même personne, même méthode, la différence maximale
## au risque de 5% qui sépare 2 résultats sur une même acquisition vaut ##
rim
# unité de NEC sur un échantillon dont la moyenne est proche de #
Mim #.#
cvrim=rim/Mim*100 ##coefficient de variation de répétabilité ##
cvrim

fidelite[2,2]<-dllrim
fidelite[2,3]<-srim2
fidelite[2,4]<-rim
fidelite[2,5]<-Mim
fidelite[2,6]<-cvrim

fidelite<-as.data.frame(fidelite)
colnames(fidelite)<-c("fidélité","dllr","srcarré","répétabilité ou reproductibilité","moyenne","CV")

```

```
write.table(fidelite, file=paste("F:/resultats_nec3d/fidelite/", paste(pointage$methode[1], pointage$grille[1]), sep=""), sep=";", dec=".", row.names=FALSE)
```

```
#####  
##### processus de validation #####
```

```
## vérifions les bases de données des populations de validation: ##
```

```
summary(nec3d_idem)  
rownames(nec3d_idem)<-nec3d_idem[,1]  
summary(nec3d_diff)  
rownames(nec3d_diff)<-nec3d_diff[,1]
```

```
## validation sur la population de validation contenant les mêmes vaches que la population de construction:
```

```
prediction_idem<-predict(stepi, nec3d_idem, se.fit=TRUE)  
nec3didem<-prediction_idem$fit  
nec3didem  
nec3did<-as.data.frame(nec3didem)  
valid_idem<-cbind(nec3d_idem, nec3did)  
valid_idem
```

```
## qualité de l'estimation pour des vaches identiques à la population de construction mais avec des  
## acquisitions différentes:
```

```
modele_valid_idem<-lm(nec~nec3didem, data=valid_idem)  
rcarre_valididem<-summary(modele_valid_idem)$adj.r.squared  
varianceerreur_valididem<-summary(modele_valid_idem)$sigma
```

```
validdiff<-matrix("", nrow=2, ncol=7)  
validdiff[1,1]<-"valididem"  
validdiff[2,1]<-"validdiff"
```

```
## calcul des différents critères pour la validation:
```

```
biais2_3id<-(mean(valid_idem$nec)-mean(valid_idem$nec3didem))^2  
NU3id<-(1-modele_valid_idem$coefficients[2])^2*var(valid_idem$nec3didem)  
LC3id<-(1-corr(valid_idem$nec, valid_idem$nec3didem)^2)*var(valid_idem$nec)  
MSE3id<-biais2_3id+NU3id+LC3id  
effetbiais_3id<-biais2_3id/MSE3id  
effetpenite_3id<-NU3id/MSE3id  
effetreidu_3id<-LC3id/MSE3id
```

```

validdiff[1,2]<-MSE3id
validdiff[1,3]<-effetbiais_3id
validdiff[1,4]<-effetpenne_3id
validdiff[1,5]<-effetresidu_3id
validdiff[1,6]<-summary(modele_valid_idem)$r.squared
validdiff[1,7]<-sd(summary(modele_valid_idem)$residuals)

```

```

##### validation sur la population de validation contenant #####
##### des vaches différents de la population de construction: #####

```

```

prediction_diff<-predict(stepi,nec3d_diff,se.fit=TRUE)
nec3ddiff<-prediction_diff$fit
nec3ddiff
nec3ddi<-as.data.frame(nec3ddiff)
valid_diff<-cbind(nec3d_diff,nec3ddi)
valid_diff

```

```

## qualité de l'estimation pour des vaches autres que celles de la population de construction mais
## avec des acquisitions différentes:

```

```

modele_valid_diff<-lm(nec~nec3ddiff,data=valid_diff)
rcarre_validdiff<-summary(modele_valid_diff)$adj.r.squared
rcarre_validdiff
varianceerreur_validdiff<-summary(modele_valid_diff)$sigma
varianceerreur_validdiff

```

```

biais2_3dif<-(mean(valid_diff$nec)-mean(valid_diff$nec3ddiff))^2
NU3dif<-(1-modele_valid_diff$coefficients[2])^2*var(valid_diff$nec3ddiff)
LC3dif<-(1-cor(valid_diff$nec,valid_diff$nec3ddiff)^2)*var(valid_diff$nec)
MSE3dif<-biais2_3dif+NU3dif+LC3dif
effetbiais_3dif<-biais2_3dif/MSE3dif
effetpenne_3dif<-NU3dif/MSE3dif
effetresidu_3dif<-LC3dif/MSE3dif
validdiff[2,2]<-MSE3dif
validdiff[2,3]<-effetbiais_3dif
validdiff[2,4]<-effetpenne_3dif
validdiff[2,5]<-effetresidu_3dif
validdiff[2,6]<-summary(modele_valid_diff)$r.squared
validdiff[2,7]<-sd(summary(modele_valid_diff)$residuals)

```

```

validdiff<-as.data.frame(validdiff)

```

```

colnames(validdiff)<-c("validation","biaisglobal","biaismoyenne","biaispente","biaisinexpliqué","rcarré","erreur
prediction")
fichiervalidation<-paste(validdiff$methode[1],validdiff$grille[1])
write.table(validdiff,file=paste("F:/resultats_nec3d/validation/",paste(valid_diff$methode[1],valid_diff$grille[1],sep=
""),sep=""),sep=";",dec=" ",row.names=FALSE)

windows()
plot(valid_diff$nec3ddiff,valid_diff$nec,type="p",main="validation de nec3d sur la population
validdiff",col=4,xlim=c(0,5),ylim=c(0,5))
abline(modele_valid_diff,col=4) abline(a=0,b=1,lty=2,col=1)
savePlot(filename=paste("F:/resultats_nec3d/validation/",paste(valid_diff$methode[1],valid_diff$grille[1],valid_diff$po
pulation[1],sep=""),sep=""),type="jpeg",device=dev.cur())

plot(valid_idem$nec3didem,valid_idem$nec,type="p",main="validation de nec3d sur la population
valididem",col=2,xlim=c(0,5),ylim=c(0,5))
abline(modele_valid_idem,col=2) ## représentation graphique du modèle abline(a=0,b=1,lty=2,col=1) ## représentation
graphique de la première bissectrice
savePlot(filename=paste("F:/resultats_nec3d/validation/",paste(valid_idem$methode[1],valid_idem$grille[1],valid_idem$po
pulation[1],sep=""),sep=""),type="jpeg",device=dev.cur())

#####
##### extraction des résidus des modèles appliqués sur les 2 populations de validation #####

## extraction des résidus pour chacune des populations de validation:

residus_validdiff<-as.data.frame(modele_valid_diff$residuals)
residus_validdiff<-as.data.frame(residus_validdiff[,1])
colnames(residus_validdiff)<-paste(valid_diff$methode[1],valid_diff$grille[1],sep="")

residus_valididem<-as.data.frame(modele_valid_idem$residuals)
residus_valididem<-as.data.frame(residus_valididem[,1])
colnames(residus_valididem)<-paste(valid_idem$methode[1],valid_idem$grille[1],sep="")

resdiff<-cbind(resdiff,residus_validdiff)
write.table(resdiff,file=paste("F:/resultats_nec3d/comparaison/", "residusdiff.csv",sep=""),sep=";",dec=" ",row.names=FA
LSE)
residem<-cbind(residem,residus_valididem)

```

```
write.table(residem,file=paste("F:/resultats_nec3d/comparaison/", "residusidem.csv", sep=""), sep=";", dec=" ", row.names=FALSE)
```

```
#####  
#####  
#####  
##### méthode 4 points pour une grille 150x150 #####
```

```
construction<-subset( quatrepoints, quatrepoints$methode=="4pointsremplissage" & quatrepoints$grille=="150150" &  
quatrepoints$pop=="cowDB" )  
nec3d_idem<-subset( quatrepoints, quatrepoints$methode=="4pointsremplissage" & quatrepoints$grille=="150150" &  
quatrepoints$population=="valididem" )  
nec3d_diff<-subset( quatrepoints, quatrepoints$methode=="4pointsremplissage" & quatrepoints$grille=="150150" &  
quatrepoints$population=="validdiff" & quatrepoints$pop=="cowTest" )
```

```
summary(construction)
```

```
##sélection du meilleur modèle cherchant à expliquer les variations observées de la NEC en fonction des coordonnées  
## des surfaces 3D dans la population de calibration dans le repère de l'ACP:
```

```

adjrcarre<-
data.frame(modele=0,étape=0,rcarré=0,rcarrévalididem=0,rcarrévaliddiff=0,erreur=0,erreurvalididem=0,erreurvaliddiff=0)

###création d'une table contenant pour chaque étape du stepwise le r2 ajusté du modèle et le r2 ajusté
## du modèle sur la population validation externe (rcarrévalididem et rcarrévaliddiff), sur les
## populations pour la répétabilité (rcarrépoints et rcarréimage) et l'étape du stepwise. Cette étape
## renseigne que pour démarrer le modèle ne contient aucune variable, adjr2=0;

for (i in 1:57)
{stepi<-step(lm(nec~1,construction),nec ~
cp1+cp2+cp3+cp4+cp5+cp6+cp7+cp8+cp9+cp10+cp11+cp12+cp13+cp14+cp15+cp16+cp17+cp18+cp19+cp20+cp21+cp22+cp23+cp24+cp25+cp2
6+cp27+cp28+cp29+cp30+cp31+cp32+cp33+cp34+cp35+cp36+cp37+cp38+cp39+cp40+cp41+cp42+cp43+cp44+cp45+cp46+cp47+cp48+cp49+cp
50+cp51+cp52+cp53+cp54+cp55+cp56+cp57,steps=i,direction="both",data=construction)
outputi<-summary(stepi)
adjrsqi<-outputi$adj.r.squared
erreur1<-sd(outputi$residuals)
## population de validation avec les individus identiques à la population de construction: ##
prediction_idem<-predict(stepi,nec3d_idem,se.fit=TRUE)
predict3didem<-as.data.frame(prediction_idem$fit)
stepvalid_idem<-cbind(nec3d_idem,predict3didem)
valididem.stepi<-lm(nec~prediction_idem$fit,data=stepvalid_idem)
adjrsqvalididem<-summary(valididem.stepi)$adj.r.squared
erreurvalidid<-sd(summary(valididem.stepi)$residuals)
## population de validation avec les individus différents de la population de construction: ##
prediction_diff<-predict(stepi,nec3d_diff,se.fit=TRUE)
predict3ddiff<-as.data.frame(prediction_diff$fit)
stepvalid_diff<-cbind(nec3d_diff,predict3ddiff)
validdiff.stepi<-lm(nec~prediction_diff$fit,data=stepvalid_diff)
adjrsqvaliddiff<-summary(validdiff.stepi)$adj.r.squared
erreurvaliddif<-sd(summary(validdiff.stepi)$residuals)
##construction d'une base de données contenant tous les R2 de tous les modèles pour chacune des populations
adjrcarrei<-
data.frame(modele=0,étape=i,rcarré=adjrsqi,rcarrévalididem=adjrsqvalididem,rcarrévaliddiff=adjrsqvaliddiff,erreur=erreu
r1,erreurvalididem=erreurvalidid,erreurvaliddiff=erreurvaliddif)
adjrcarre<-rbind(adjrcarre,adjrcarrei) ## fusionne l'ancienne table r2 avec la nouvelle ligne calculée
## pour l'étape i
}

### représentation graphique de l'évolution des modèles en fonction du nombre de composantes dans
## le modèle, r2=f(i):
modelebest<-stepi$call ## extrait le modèle complet du step
modelebest

```

```

summary(stepi)
modell<-gsub("nec ~ ","",modelebest, fixed=TRUE)
  ## remplacement exactement les caractères "nec ~ " du modèle par un vide "" avec l'option fixed=TRUE
  ## qui permet de remplacer exactement les caractères
modell
model<-gsub("+","",modell[2],fixed=TRUE)
model
composante<-strsplit(model," ") ## coupe la chaîne de caractères au niveau des " "
composante<-noquote(composante[[1]])## retire les trêmas de la chaîne de caractères
composante
composante<-as.matrix(as.factor(composante))
composante
dim(composante)
cp<-as.data.frame((as.matrix(composante))) ## transformation en data.frame
dim(cp) ## donne le nombre de cp du meilleur modèle
adjrcarre
listemodele<-cbind(listemodele,cp)

write.table(adjrcarre,file=paste("F:/resultats_nec3d/calibration/",paste(construction$methode[1],construction$grille[1]
),sep=""),sep=";",dec=".",append=FALSE,quote=TRUE)

### traçons les graphes pour voir l'évolution des modèles au fur et à mesure de l'avancement du stepwise: windows() ####
ouvre la console affichant les graphes sous R ## barplot(adjrcarre$rcarré,names.arg=adjrcarre$étape,main="évolution de
la qualité des modèles du stepwise",xlab="nombre d'étapes dans la sélection par stepwise",ylab="R-carré
ajusté",ylim=c(0,1))
savePlot(filename=paste("F:/resultats_nec3d/calibration/",paste(construction$methode[1],construction$grille[1],construc
tion$population[1],construction$pop[1]),sep=""),type="jpeg",device=dev.cur())

### enregistre le graphe sous format .jpeg sur mon disque dur externe (F:). La commande path.expand
## renseigne le chemin d'accès où R doit enregistrer le document! ###
barplot(adjrcarre$rcarrévalididem,names.arg=adjrcarre$étape,main=paste("évolution de la qualité des modèles du
stepwise","appliqués sur la population de validation","avec les vaches identiques à la population de
construction",sep="\n"),xlab="nombre d'étapes dans la sélection par stepwise",ylab="R-carré ajusté",ylim=c(0,1))
savePlot(filename=paste("F:/resultats_nec3d/calibration/",paste(nec3d_idem$methode[1],nec3d_idem$grille[1],nec3d_idem$pop
ulation[1]),sep=""),type="jpeg",device=dev.cur())
barplot(adjrcarre$rcarrévaliddiff,names.arg=adjrcarre$étape,main=paste("évolution de la qualité des modèles du
stepwise","appliqués sur la population de validation", "avec les vaches différentes de la population de
construction",sep="\n"), xlab="nombre d'étapes dans la sélection par stepwise",ylab="R-carré ajusté",ylim=c(0,1))
savePlot(filename=paste("F:/resultats_nec3d/calibration/",paste(nec3d_diff$methode[1],nec3d_diff$grille[1],nec3d_diff$pop
ulation[1],nec3d_diff$pop[1]),sep=""),type="jpeg",device=dev.cur())

## graphe pour les erreurs de prédiction:

```



```

barplot(adjrcarre$erreur,names.arg=adjrcarre$étape,main="évolution de la qualité des modèles du stepwise",xlab="nombre
d'étapes dans la sélection par stepwise",ylab="erreurde prediction")
savePlot(filename=paste("F:/resultats_nec3d/calibration/","erreur",paste(construction$methode[1],construction$grille[1]
,construction$population[1],construction$pop[1]),sep=""),type="jpeg",device=dev.cur())

### enregistre le graphe sous format .jpeg sur mon disque dur externe (F:). La commande path.expand
## renseigne le chemin d'accès où R doit enregistrer le document! ###
barplot(adjrcarre$erreurvalididem,names.arg=adjrcarre$étape,main=paste("évolution de la qualité des modèles du
stepwise","appliqués sur la population de validation","avec les vaches identiques à la population de
construction",sep="\n"),xlab="nombre d'étapes dans la sélection par stepwise",ylab="erreur de
prediction",ylim=c(0,0.8))
savePlot(filename=paste("F:/resultats_nec3d/calibration/","erreur",paste(nec3d_idem$methode[1],nec3d_idem$grille[1],nec
3d_idem$population[1]),sep=""),type="jpeg",device=dev.cur())
barplot(adjrcarre$erreurvaliddiff,names.arg=adjrcarre$étape,main=paste("évolution de la qualité des modèles du
stepwise","appliqués sur la population de validation", "avec les vaches différentes de la population de
construction",sep="\n"), xlab="nombre d'étapes dans la sélection par stepwise",ylab="erreur de
prediction",ylim=c(0,0.8))
savePlot(filename=paste("F:/resultats_nec3d/calibration/","erreur",paste(nec3d_diff$methode[1],nec3d_diff$grille[1],nec
3d_diff$population[1],nec3d_diff$pop[1]),sep=""),type="jpeg",device=dev.cur())

windows()
plot(stepi$fitted.values,construction$nec,type="p",main=paste("qualité du modèle de
calibration",paste(construction$methode[1],construction$grille[1],sep=""),sep=""),xlab="nec prédite",ylab="nec
observées",xlim=c(0,5),ylim=c(0,5))
abline(lm(construction$nec~stepi$fitted.values)) abline(a=0,b=1,col="red",lty=2)
savePlot(filename=paste("F:/resultats_nec3d/calibration/","modelecalib",paste(construction$methode[1],construction$gril
le[1],sep=""),sep=""),type="jpeg",device=dev.cur())

#####
##### processus de validation #####

## vérifions les bases de données des populations de validation: ##

summary(nec3d_idem)
rownames(nec3d_idem)<-nec3d_idem[,1]
summary(nec3d_diff)
rownames(nec3d_diff)<-nec3d_diff[,1]

## validation sur la population de validation contenant les mêmes vaches que la population de construction:
prediction_idem<-predict(stepi,nec3d_idem,se.fit=TRUE)

```

```

nec3didem<-prediction_idem$fit
nec3didem
nec3did<-as.data.frame(nec3didem)
valid_idem<-cbind(nec3d_idem,nec3did)
valid_idem

## qualité de l'estimation pour des vaches identiques à la population de construction mais avec des
## acquisitions différentes:

modele_valid_idem<-lm(nec~nec3didem,data=valid_idem)
rcarre_valididem<-summary(modele_valid_idem)$adj.r.squared
varianceerreur_valididem<-summary(modele_valid_idem)$sigma

validdiff<-matrix("",nrow=2,ncol=7)
validdiff[1,1]<-"valididem"
validdiff[2,1]<-"validdiff"

## calcul des différents critères pour la validation:
biais2_3id<-(mean(valid_idem$nec)-mean(valid_idem$nec3didem))^2
NU3id<-(1-modele_valid_idem$coefficients[2])^2*var(valid_idem$nec3didem)
LC3id<-(1-cor(valid_idem$nec,valid_idem$nec3didem)^2)*var(valid_idem$nec)
MSE3id<-biais2_3id+NU3id+LC3id
effetbiais_3id<-biais2_3id/MSE3id
effetpenete_3id<-NU3id/MSE3id
effetresidu_3id<-LC3id/MSE3id
validdiff[1,2]<-MSE3id
validdiff[1,3]<-effetbiais_3id
validdiff[1,4]<-effetpenete_3id
validdiff[1,5]<-effetresidu_3id
validdiff[1,6]<-summary(modele_valid_idem)$r.squared
validdiff[1,7]<-sd(summary(modele_valid_idem)$residuals)

##### validation sur la population de validation contenant #####
##### des vaches différentes de la population de construction: #####

prediction_diff<-predict(stepi,nec3d_diff,se.fit=TRUE)
nec3ddiff<-prediction_diff$fit
nec3ddiff
nec3ddi<-as.data.frame(nec3ddiff)
valid_diff<-cbind(nec3d_diff,nec3ddi)
valid_diff

## qualité de l'estimation pour des vaches autres que celles de la population de construction mais avec

```

```
## des acquisitions différentes:
```

```
modele_valid_diff<-lm(nec~nec3ddiff,data=valid_diff)  
rcarre_validdiff<-summary(modele_valid_diff)$adj.r.squared  
rcarre_validdiff  
varianceerreur_validdiff<-summary(modele_valid_diff)$sigma  
varianceerreur_validdiff
```

```
biais2_3dif<-(mean(valid_diff$nec)-mean(valid_diff$nec3ddiff))^2  
NU3dif<-(1-modele_valid_diff$coefficients[2])^2*var(valid_diff$nec3ddiff)  
LC3dif<-(1-cor(valid_diff$nec,valid_diff$nec3ddiff)^2)*var(valid_diff$nec)  
MSE3dif<-biais2_3dif+NU3dif+LC3dif  
effetbiais_3dif<-biais2_3dif/MSE3dif  
effetpenste_3dif<-NU3dif/MSE3dif  
effetresidu_3dif<-LC3dif/MSE3dif  
validdiff[2,2]<-MSE3dif  
validdiff[2,3]<-effetbiais_3dif  
validdiff[2,4]<-effetpenste_3dif  
validdiff[2,5]<-effetresidu_3dif  
validdiff[2,6]<-summary(modele_valid_diff)$r.squared  
validdiff[2,7]<-sd(summary(modele_valid_diff)$residuals)
```

```
validdiff<-as.data.frame(validdiff)  
colnames(validdiff)<-c("validation","biaisglobal","biaismoyenne","biaispenste","biaisinexpliqué","rcarré","erreur  
prediction")  
fichiervalidation<-paste(validdiff$methode[1],validdiff$grille[1])  
write.table(validdiff,file=paste("F:/resultats_nec3d/validation/",paste(valid_diff$methode[1],valid_diff$grille[1],sep=  
""),sep=""),sep=";",dec=" ",row.names=FALSE)
```

```
windows()  
plot(valid_diff$nec3ddiff,valid_diff$nec,type="p",main="validation de nec3d sur la population  
validdiff",col=4,xlim=c(0,5),ylim=c(0,5))  
abline(modele_valid_diff,col=4)  
abline(a=0,b=1,lty=2,col=1)  
savePlot(filename=paste("F:/resultats_nec3d/validation/",paste(valid_diff$methode[1],valid_diff$grille[1],valid_diff$po  
pulation[1],sep=""),sep=""),type="jpeg",device=dev.cur())
```

```
plot(valid_idem$nec3didem,valid_idem$nec,type="p",main="validation de nec3d sur la population  
valididem",col=2,xlim=c(0,5),ylim=c(0,5))  
abline(modele_valid_idem,col=2) ## représentation graphique du modèle  
abline(a=0,b=1,lty=2,col=1) ## représentation graphique de la première bissectrice
```

```
savePlot(filename=paste("F:/resultats_nec3d/validation/",paste(valid_idem$methode[1],valid_idem$grille[1],valid_idem$population[1],sep=""),sep=""),type="jpeg",device=dev.cur())
```

```
#####  
##### extraction des résidus des modèles appliqués sur les 2 populations de validation #####
```

```
## extraction des résidus pour chacune des populations de validation:
```

```
residus_validdiff<-as.data.frame(modele_valid_diff$residuals)  
residus_validdiff<-as.data.frame(residus_validdiff[,1])  
colnames(residus_validdiff)<-paste(valid_diff$methode[1],valid_diff$grille[1],sep="")
```

```
residus_valididem<-as.data.frame(modele_valid_idem$residuals)  
residus_valididem<-as.data.frame(residus_valididem[,1])  
colnames(residus_valididem)<-paste(valid_idem$methode[1],valid_idem$grille[1],sep="")
```

```
resdiff<-cbind(resdiff,residus_validdiff)  
write.table(resdiff,file=paste("F:/resultats_nec3d/comparaison/","residusdiff.csv",sep=""),sep=";",dec=" ",row.names=FALSE)  
residem<-cbind(residem,residus_valididem)  
write.table(residem,file=paste("F:/resultats_nec3d/comparaison/","residusidem.csv",sep=""),sep=";",dec=" ",row.names=FALSE)
```

```
##### FIN DE REPETITION DE LA PROCEDURE #####  
#####  
#####  
#####
```

```
##### export du tableau contenant la liste des 4 modèles sélectionnés par stepwise:
```

```
write.table(listmodele,file=paste("F:/resultats_nec3d/calibration/","modeles.csv",sep=""),sep=";",append=FALSE,quote=TRUE,row.names=FALSE)
```

```
#####  
#####  
##### comparaison des méthodes et des grilles #####
```

```

## quelle est la corrélation entre les résidus des 2 modèles?
windows()
cordiffpts<-cor.test(resdiff[,2],resdiff[,5]) ## calcul du coefficient de corrélation entre les résidus
                                             ## des modèles de la méthode "3points"
                                             ## avec ceux de la méthode "4points" pour la population "validdiff"

cordiffpts
plot(resdiff[,2],resdiff[,5],type="n",xlim=c(-0.6,1),ylim=c(-0.6,1),main=paste("comparaison des 2 méthodes de
pointage","sur la population de validation validdiff",sep="\n"),xlab="résidus pointage fesses",ylab="résidus pointage
attache de la queue")
    ## représentation graphique des résidus de la méthode "4points" en fonction des résidus de
    ## la méthode "3points"
abline(frect(resdiff[,2],resdiff[,5]),col="blue")
    ## représente la droite de régression linéaire ajustée sur les données avec la méthode
    ## des moindres rectangles, à l'aide de la fonction frect(x,y) introduite au début du script
abline(a=0,b=1,lty=2) ## représentation graphique de la lère bissectrice text(resdiff[,2],resdiff[,5],resdiff[,1]) ##
affiche les identifiant des vaches sur le graphe
savePlot(filename=paste("F:/resultats_nec3d/comparaison/",paste("methode",valid_diff$population[1],sep=""),sep=""),type
="jpeg",device=dev.cur())
    ## exporte le graphique sur le disque dur

## répétition de la même procédure pour la population "valididem":
coridpts<-cor.test(residem[,2],residem[,5])$estimate
coridpts
plot(residem[,2],residem[,5],type="n",xlim=c(-0.6,1),ylim=c(-0.6,1),main=paste("comparaison des 2 méthodes de
pointage","sur la population de validation valididem",sep="\n"),ylab="résidus pointage attache de la
queue",xlab="résidus pointage fesses")
abline(frect(residem[,2],residem[,5]),col="blue")
abline(a=0,b=1,col="black",lty=2) text(residem[,2],residem[,5],residem[,1])
savePlot(filename=paste("F:/resultats_nec3d/comparaison/",paste("methode",valid_idem$population[1],sep=""),sep=""),type
="jpeg",device=dev.cur())

## quelle est la corrélation entre les résidus 3 grilles?
## répétition de la procédure pour la comparaison entre la résolution "150x150" et "100x100"
## pour la population "validdiff":
cordiff150100<-cor.test(resdiff[,2],resdiff[,3])$estimate
cordiff150100
plot(resdiff[,2],resdiff[,3],type="n",xlim=c(-0.6,1),ylim=c(-0.6,1),main=paste("comparaison des 2 grilles: 150150 et
100100","sur la population de validation validdiff",sep="\n"),ylab="résidus grille 100100",xlab="résidus residus grille
150150")
abline(frect(resdiff[,2],resdiff[,3]),col="red")

```

```

abline(a=0,b=1,col="black") text(resdiff[,2],resdiff[,3],resdiff[,1])
savePlot(filename=paste("F:/resultats_nec3d/comparaison/", "grille150100diff", sep=""), type="jpeg", device=dev.cur())

## pour la population "valididem":
corid150100<-cor.test(residem[,2],residem[,3])$estimate
corid150100
plot(residem[,2],residem[,3],type="n",xlim=c(-0.6,1),ylim=c(-0.6,1),main=paste("comparaison des 2 grilles: 150150 et
100100", "sur la population de validation valididem", sep="\n"),ylab="résidus grille 100100",xlab="résidus residus grille
150150")
abline(frect(residem[,2],residem[,3]),col="red")
abline(a=0,b=1,col="black")
text(residem[,2],residem[,3],residem[,1])
savePlot(filename=paste("F:/resultats_nec3d/comparaison/", "grille150100idem", sep=""), type="jpeg", device=dev.cur())

## répétition de la procédure pour la comparaison entre la résolution "150x150" et "50x50"
## pour la population "validdiff":
cordiff15050<-cor.test(resdiff[,2],resdiff[,4])$estimate
cordiff15050
plot(resdiff[,2],resdiff[,4],type="n",xlim=c(-0.6,1),ylim=c(-0.6,1),main=paste("comparaison des 2 grilles: 150150 et
5050", "sur la population de validation validdiff", sep="\n"),ylab="résidus grille 5050",xlab="résidus residus grille
150150")
abline(frect(resdiff[,2],resdiff[,4]),col="red")
abline(a=0,b=1,col="black")
text(resdiff[,2],resdiff[,4],resdiff[,1])
savePlot(filename=paste("F:/resultats_nec3d/comparaison/", "grille15050diff", sep=""), type="jpeg", device=dev.cur())

## pour la population "valididem":
corid15050<-cor.test(residem[,2],residem[,4])$estimate
corid15050
plot(residem[,2],residem[,4],type="n",xlim=c(-0.6,1),ylim=c(-0.6,1),main=paste("comparaison des 2 grilles: 150150 et
5050", "sur la population de validation valididem", sep="\n"),ylab="résidus grille 5050",xlab="résidus residus grille
150150")
abline(frect(residem[,2],residem[,4]),col="red")
abline(a=0,b=1,col="black")
text(residem[,2],residem[,4],residem[,1])
savePlot(filename=paste("F:/resultats_nec3d/comparaison/", "grille15050idem", sep=""), type="jpeg", device=dev.cur())

## répétition de la procédure pour la comparaison entre la résolution "50x50" et "100x100"
## pour la population "validdiff":
cordiff50100<-cor.test(resdiff[,4],resdiff[,3])$estimate
cordiff50100

```

```
plot(resdiff[,4],resdiff[,3],type="n",xlim=c(-0.6,1),ylim=c(-0.6,1),main=paste("comparaison des 2 grilles: 5050 et
100100","sur la population de validation validdiff",sep="\n"),ylab="résidus grille 100100",xlab="résidus residus grille
5050")
abline(frect(resdiff[,4],resdiff[,3]),col="red")
abline(a=0,b=1,col="black")
text(resdiff[,4],resdiff[,3],resdiff[,1])
savePlot(filename=paste("F:/resultats_nec3d/comparaison/","grille50100diff",sep=""),type="jpeg",device=dev.cur())
```

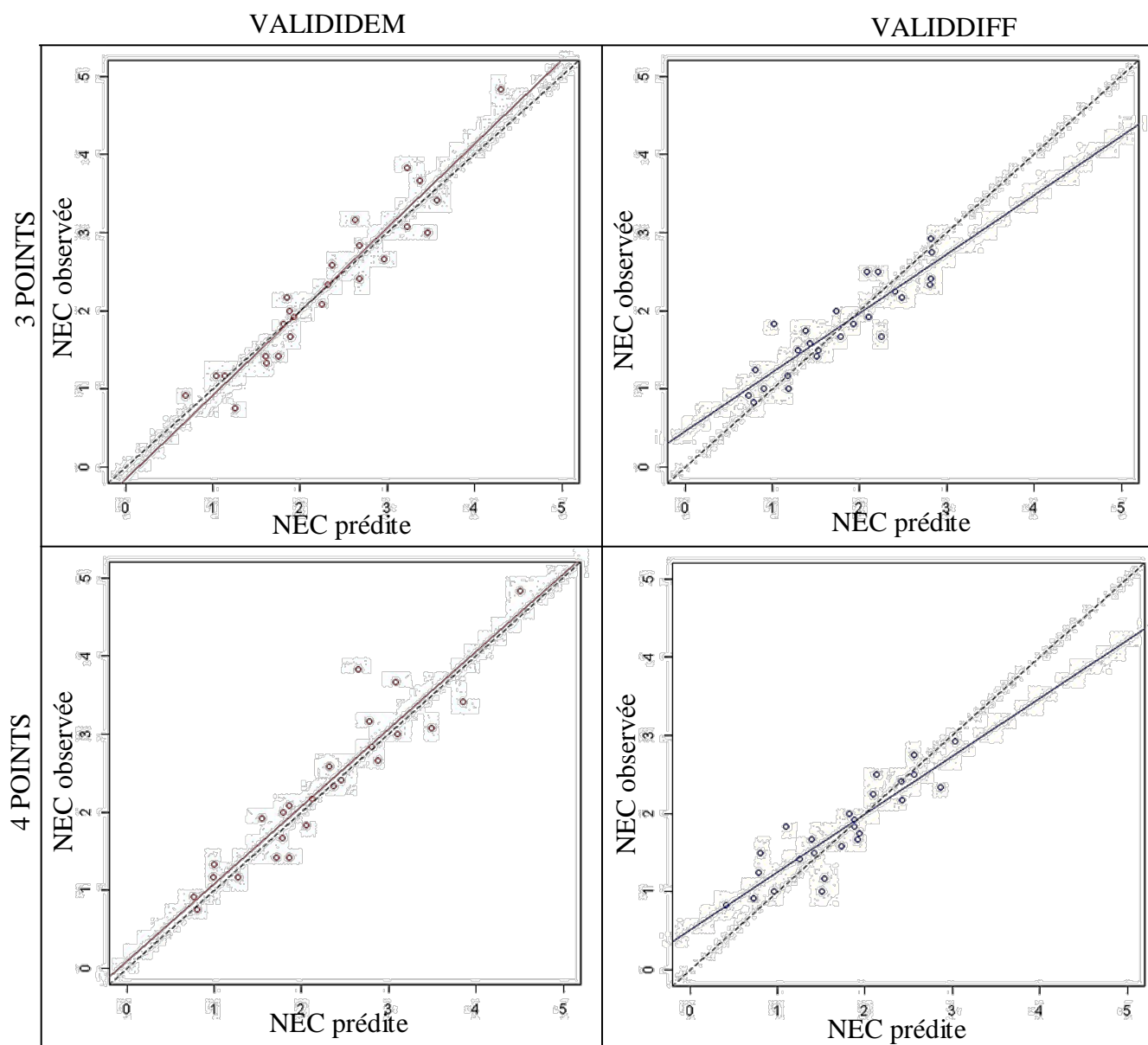
```
## pour la population "valididem":
corid50100<-cor.test(residem[,4],residem[,3])$estimate
corid50100
plot(residem[,4],residem[,3],type="n",xlim=c(-0.6,1),ylim=c(-0.6,1),main=paste("comparaison des 2 grilles: 5050 et
100100","sur la population de validation valididem",sep="\n"),ylab="résidus grille 100100",xlab="résidus residus grille
5050") text(residem[,4],residem[,3],residem[,1])
abline(frect(residem[,4],residem[,3]),col="red")
abline(a=0,b=1,col="black")
savePlot(filename=paste("F:/resultats_nec3d/comparaison/","grille50100idem",sep=""),type="jpeg",device=dev.cur())
```

```
## création d'une table contenant tous les coefficients de corrélation:
correlations<-rbind(cordiffpts,coridpts,cordiff150100,corid150100,cordiff15050,corid15050,cordiff50100,corid50100)
row.names(correlations)<-
c("cordiffpts","coridpts","cordiff150100","corid150100","cordiff15050","corid15050","cordiff50100","corid50100")
      ## identification des lignes
colnames(correlations)<-c("correlations") ## identification des colonnes
```

```
write.table(correlations,file=paste("F:/resultats_nec3d/comparaison/","correlations.csv",sep=""),sep=";",dec=",")
      ## exporte la table sur le disque dur
```

```
#####
#####
#####          FIN DE PROCEDURE          #####
#####
#####
```

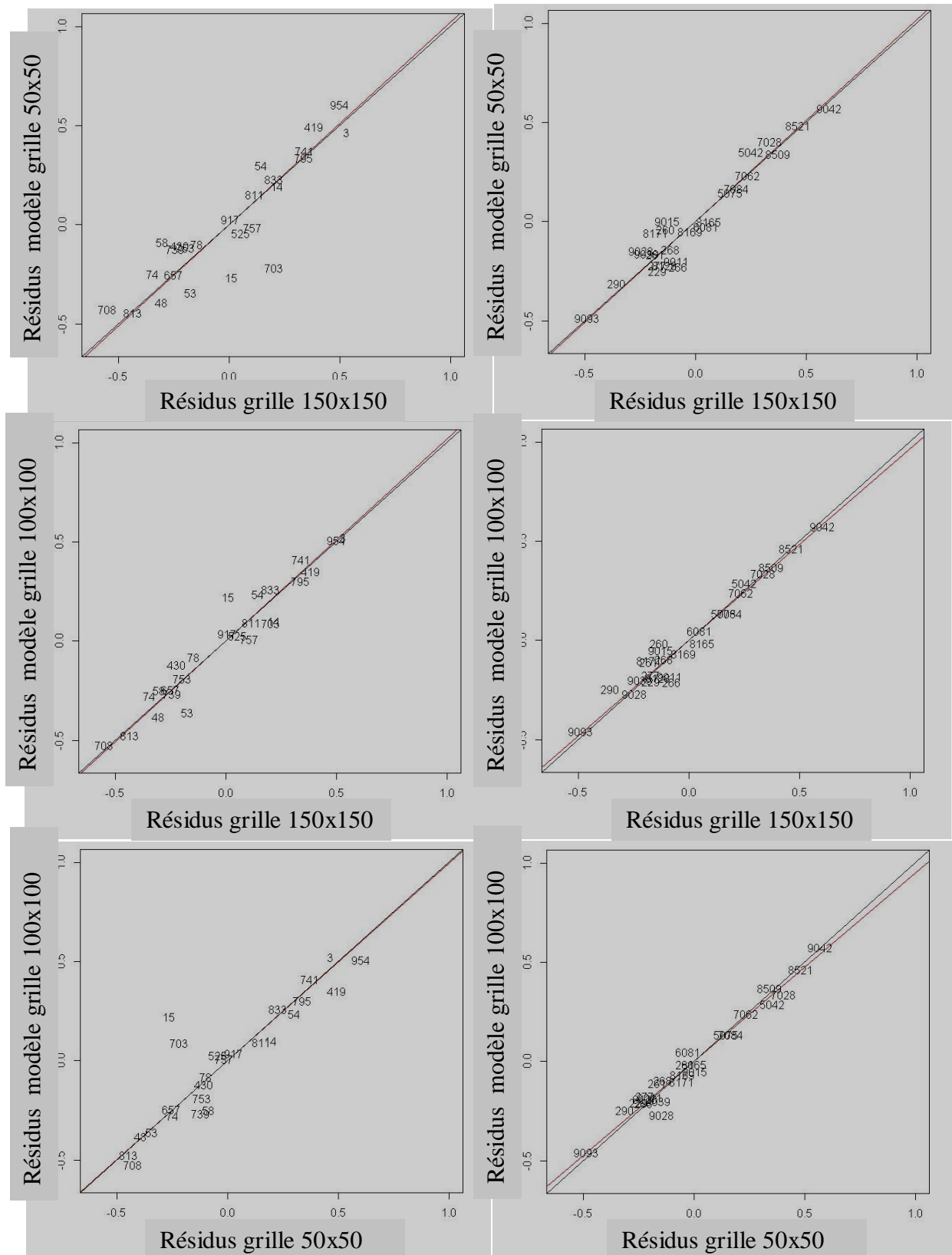
Annexe IV : Comparaison de la qualité des modèles d'estimation de la NEC pour les 2 méthodes de pointage (3points, 4 points) et pour les 2 populations de validation externe (valididem, validdiff) par régression des NEC observées en fonction des NEC prédites



--- Première bissectrice

— Droite de régression ajustée sur les données par minimisation des moindres rectangles

Annexe V : Comparaison de la qualité des modèles d'estimation de la NEC pour les 3 résolutions (150x150, 100x100, 50x50) pour les 2 population « valididem » (à gauche) et « validdiff » (à droite)



- Première bissectrice
- Droite de régression ajustée sur les données par minimisation des moindres rectangles



Diplôme : Diplôme D'Ingénieur de l'Institut Supérieur des Sciences Agronomiques, Agroalimentaires, Horticoles et du Paysage

Spécialité : Sciences Animales pour l'Élevage de Demain (SAED)

Spécialisation / option : parcours professionnel

Enseignant référent : Sandrine LAGARRIGUE

Auteur(s) : Amélie FISCHER

Date de naissance* : 17/12/1990

Nb pages :24 p. Annexe(s) : 5 (54 p.)

Année de soutenance : 2013

Organisme d'accueil : INRA –UMR PEGASE

Adresse : Domaine de la Prise

35590 Saint-Gilles

Maître de stage : Philippe FAVERDIN

Titre français : Utilisation de l'imagerie 3D pour estimer la note d'état corporel des vaches laitières

Titre anglais : 3D imaging for estimating body condition scores in dairy cattle

Résumé :

L'état corporel affecte les performances de reproduction et de santé des vaches laitières. En élevage, des experts l'évaluent couramment avec la note d'état corporel (NEC) parce qu'elle est peu chère mais subjective et peu sensible. Le projet vise à développer et à valider une NEC3D qui estime la NEC à partir de surfaces 3D allant des pointes de fesses aux pointes de hanches des vaches. Une analyse en composantes principales a été appliquée au jeu de calibration contenant 57 surfaces. Les composantes principales ont été incrémentées dans un modèle de régression linéaire multiple pour estimer la NEC. La normalisation des surfaces a nécessité un pointage manuel de quatre points anatomiques. L'influence de deux types de pointage et de la résolution des surfaces 3D sur la qualité de la méthode a été analysée. Deux validations externes ont été réalisées : une sur des vaches du jeu de calibration mais à un stade de lactation différent (valididem) et une sur des vaches externes au jeu de calibration (validdiff). Les deux méthodes de pointages sont similaires en termes de calibration ($R^2=1$) mais diffèrent un peu en termes de qualité de validation (RMSE = 0.34 vs. 0.32). La NEC3D est 2.8 fois plus reproductible que la NEC ($\sigma = 0.1$ vs. 0.28). Comme l'erreur de reproductibilité tient compte de l'erreur de répétabilité, elle-même associée à l'erreur liée au pointage manuelle, l'automatisation du pointage se présente comme une étape indispensable pour améliorer la répétabilité et par conséquent la reproductibilité. Une baisse de résolution, bien que diminuant la reproductibilité, conserve la même erreur de prédiction. Ceci motive le développement d'un système d'acquisition plus simple, associé à des résolutions plus faible. L'imagerie 3D ressort comme un outil encourageant pour développer le phénotypage de l'état corporel.

Abstract :

Body condition is known to affect reproduction and health in dairy cows. On-farm it is usually measured with the body condition score (BCS) which is not expensive but remains subjective and not very sensitive. The aim of the current work was to develop and to validate a method, nec3D, working on estimating BCS with 3D pictures of dairy cattle's back, from the pins to the hooks. A 57 3D-shapes dataset, transformed with a principal component analysis, was built for calibration. The principal components were performed on BCS with multiple linear regressions. Four anatomical points had to be identified manually to normalise the pictures. Influence of two different ways of points' identification and of the picture's resolution on methods quality was analysed. Moreover, external validation was evaluated on two additional datasets: one with cows used for calibration, but with different stages in milking (valididem) and one with cows not used for calibration (validdiff). Both ways of points' identification had quite good results in terms of calibration ($R^2 = 1$) and differed slightly on validation quality (RMSE= 0.34 vs. 0.32). Nec3d was 2.8 times more reproducible than BCS ($\sigma = 0.1$ vs. 0.28). The lowest resolution implied a loss of reproducibility but did not increase the error of prediction. A simplified acquisition system, implying low resolution, could therefore be developed. As error of reproducibility incorporates error due to the manual point's identification, automation of points' identification would improve repeatability and consequently reproducibility. Assessing body condition thanks to 3D pictures appears as a promising tool which can improve phenotyping of this trait.

Mots-clés : imagerie 3D, état corporel, vache laitière, élevage de précision

Keywords: 3D imaging, body condition, dairy cows, precision livestock farming