
Détection de ruptures dans la trajectoire des données de production et de comportement chez la vache laitière

Mise en relation avec les événements sanitaires, de reproduction et d'alimentation du troupeau

- 2^{ème} édition -

Rédigé par Amandine ROUILLARD

Master 2 éthologie - Comportement animal et humain



Université de Rennes

Stage réalisé du 15 Janvier 2024 au 14 Juin 2024,
à l'INRAE - UMR 1348 PEGASE, équipe Sysmo (Saint-Gilles)

Soutenance orale le 20 Juin 2024, à Rennes

Encadrée par

Charlotte GAILLARD et

Anne BOUDON, chercheuses à l'INRAE, Saint Gilles - UMR 1348 PEGASE,

Tom ROHMER, chercheur à l'INRAE, Toulouse - UMR 1388 GenPhySE.

Référente universitaire

Sophie LUMINEAU,

enseignante-chercheuse à l'université de Rennes

Ce travail a bénéficié d'une aide de l'État gérée par l'Agence Nationale de la Recherche au titre du programme d'Investissements d'Avenir portant la référence ANR-16-CONV-0004

Remerciements

Je remercie tout d'abord mes encadrants, Anne Boudon, Tom Rohmer et Charlotte Gaillard, pour l'opportunité qu'ils m'ont offerte avec ce stage ainsi que pour leur accompagnement et leur pédagogie durant ces 5 mois. Je sors de ce stage avec de toutes nouvelles compétences sur R...

Merci également à toutes les personnes que j'ai eues le plaisir de rencontrer sur le site de Saint-Gilles ou avec lesquelles j'ai pu échanger par mail : aux membres des équipes de Pegase, BIENETRE et SYSMO, pour l'ouverture pluridisciplinaire qu'ils m'ont apporté, à Ophélie pour l'appui technique, à Myriam pour son soutien administratif, sa disponibilité et sa bienveillance, à Nina et Wilfried pour leur bonne humeur communicative partagée chaque midi. Merci à Anna et aux autres stagiaires, en particulier Mélodie, Adèle, Florie-Anne et Rose-Marie, pour leurs conseils et leur appui durant les moments de doutes.

Merci à #Digitag pour avoir rendu ce projet possible ainsi que pour la communauté qu'il rassemble depuis quelques années. Merci pour l'organisation de la Digitagora de Montpellier, cela a été une étape enrichissante de mon stage.

Et bien entendu, merci aux professeurs du Master CAH de Rennes pour leur investissement dans leurs enseignements, leur partage de connaissances et d'expériences. Ces deux années ont été riches et passionnantes grâce à vous.

Enfin, je remercie mes ami(e)s et ma famille pour leur soutien infaillible et tous les bons moments passés ensemble malgré les perturbations environnantes. Ces derniers mois ont été difficiles sur bien des aspects, mais votre présence s'est avérée être une aide précieuse. Ce rapport marque la fin d'une étape et le début d'une autre. L'avenir nous dira quelles nouvelles aventures elle nous réserve...

Abraham, j'espère partager ces aventures avec toi. Ton ambition et ton courage m'inspirent et me poussent à me dépasser chaque fois un peu plus. Tu me donnes la force d'aller chercher en moi des ressources que je ne soupçonnais pas. Être aux côtés d'une personne aussi passionnante et attentionnée que toi est une chance ! 5.3..8.2.4...6.3..

Sommaire

| | |
|---|-----------|
| Remerciements | 2 |
| Sommaire | 3 |
| Introduction | 1 |
| Matériel et méthodes | 5 |
| 1. Note éthique..... | 5 |
| 2. Animaux et hébergement..... | 6 |
| 3. Procédure expérimentale..... | 6 |
| 4. Mesures zootechniques et comportementales..... | 7 |
| 5. Analyses statistiques..... | 8 |
| 5.1 <i>Nettoyage des données et analyse descriptive</i> | 8 |
| 5.2 <i>Méthodes de lissage : FDA et UpDown</i> | 9 |
| 5.3 <i>Calcul de sensibilité et spécificité</i> | 11 |
| Résultats | 14 |
| 1. Effet de la stratégie d'alimentation sur la production laitière, l'ingestion et le temps passé debout..... | 14 |
| 2. Choix du niveau de lissage..... | 15 |
| 3. Détection des perturbations..... | 17 |
| 4. Calculs de sensibilité et spécificité des méthodes..... | 18 |
| Discussion | 20 |
| 1. Fonctionnement, paramétrage et application des méthodes..... | 21 |
| 2. Robustesse des animaux et choix des variables étudiées..... | 23 |
| 3. Fiabilité des données collectées par les capteurs et les automates..... | 24 |
| 4. Constitution du gold standard..... | 25 |
| Conclusion | 26 |
| Bibliographie | 27 |
| Annexe | 30 |
| Caractérisation des perturbations selon le type d'événement..... | 30 |
| Résumé | 33 |
| Abstract | 33 |

Introduction

Le respect du bien-être des animaux est un enjeu qui concerne de nombreux domaines (*e.g.*, expérimentation animale, élevage) et qui ne cesse de soulever de nouvelles questions éthiques et scientifiques. Les attentes des citoyens européens, comme dans d'autres régions du monde, poussent les professionnels et les législateurs à adapter leurs pratiques et le cadre légal autour de la détention et l'utilisation des animaux, en particulier dans le domaine de l'élevage (European Commission, 2023 ; Sinclair *et al.*, 2022 ; ANSES, 2018).

Cette prise de conscience a été initiée dans le milieu scientifique au Royaume-Uni en 1965 par la constitution du Comité Brambell (Reimert *et al.*, 2023). Ce comité avait pour objectif d'établir un état des lieux des conditions d'élevage dans le pays et du bien-être des animaux de rente. Ce projet a abouti à la création du Farm Animal Welfare Council (FAWC) qui a proposé une première définition de la notion de bien-être animal avec le concept des 5 libertés : absence de faim, de soif et de malnutrition, absence de peur et de détresse, absence de stress physique et/ou thermique, absence de douleur, de lésions et de maladie, liberté d'expression d'un comportement normal de l'espèce (FAWC, 2009). De multiples autres définitions ont ensuite été proposées depuis les débuts des recherches dans ce domaine. En 2018, l'Agence nationale de sécurité sanitaire de l'alimentation, de l'environnement et du travail (ANSES) pose un cadre théorique autour du concept de bien-être animal en proposant une définition conceptuelle : "le bien-être d'un animal est l'état mental et physique positif lié à la satisfaction de ses besoins physiologiques et comportementaux, ainsi que de ses attentes. Cet état varie en fonction de la perception de la situation par l'animal" (ANSES, 2018). De nouvelles notions tendent à s'ajouter à cette définition. D'une part, le développement du concept de "bien-être positif" qui milite pour ne plus seulement limiter les expériences négatives mais aussi favoriser les expériences positives procurant par exemple du plaisir ou du confort à l'animal (Mellor *et al.*, 2017). Et d'autre part, l'accumulation des expériences et l'équilibre entre les événements positifs et négatifs tout au long de la vie de l'animal semblent nécessaires à intégrer dans l'évaluation de son bien-être (Reimert *et al.*, 2023).

Dans le milieu de l'élevage, il existe de nombreux protocoles d'évaluation du bien-être animal dont les indicateurs sont établis selon les besoins et les caractéristiques des espèces concernées. Le protocole Welfare Quality®, largement utilisé en Europe, repose sur le modèle des 5 domaines inspirés des 5 libertés précédemment décrites (*i.e.*, alimentation, santé, environnement, comportements, état mental, Mellor *et al.*, 2017 ; FAWC, 2009). Ce protocole permet une évaluation holistique à partir d'indicateurs validés scientifiquement (*e.g.*, accès aux

ressources mises à disposition par l'éleveur, score d'état corporel, réaction face à l'approche de l'humain, comportements sociaux). Cependant, l'utilisation de ce protocole est coûteuse et chronophage ce qui le rend difficilement applicable en routine dans les fermes commerciales malgré sa fiabilité. Par exemple, pour évaluer le bien-être d'un troupeau de 25 vaches laitières, l'éleveur doit y consacrer 4 heures de son temps (Blokhuis *et al.*, 2010). De plus, cette évaluation se fait ponctuellement et donne une image à un *instant t* de l'état de bien-être des animaux sans prendre en compte l'accumulation des expériences, ni d'éventuelles perturbations pouvant s'être produites avant les observations (*e.g.*, nouvel allotement, événement stressant).

En effet, le bien-être des animaux peut être impacté par des perturbations environnementales liées aux contraintes et aux changements inhérents aux conditions d'élevage. Dans cette étude, nous définissons une perturbation comme tous changements à court terme dans l'environnement de l'animal ayant un impact sur ses performances (Taghipoor *et al.*, 2023). Les causes de ces perturbations peuvent être de différentes natures (*e.g.*, pathologies, alimentation, facteurs climatiques, changement d'environnement, interaction avec l'Homme ou d'autres animaux, Ben Abdelkrim *et al.*, 2021 ; Nguyen-Ba *et al.*, 2020). L'effet de ces événements sur les animaux et leur bien-être dépend des caractéristiques de l'événement perturbateur (intensité, durée, fréquence) et des capacités de réponse de l'individu défini par sa résistance (*i.e.*, capacité de l'animal à minimiser l'impact de la perturbation) et sa résilience (*i.e.*, capacité de l'animal à revenir à son état initial, Le, 2022 ; Nguyen-Ba *et al.*, 2020 ; Sauvart & Martin, 2010). Il existe différentes méthodes permettant de détecter une perturbation, puis de quantifier la réponse de l'animal, comme l'utilisation de courbes théoriques (*i.e.*, trajectoire des données sans perturbations, par exemple, des courbes paramétriques de type Wilmlink pour la lactation, Sitkowska *et al.*, 2020) ou de courbes de lissage (*e.g.*, non-paramétriques de type polynômes locaux, splines ou à noyaux, Taghipoor *et al.*, 2023 ; David *et al.*, 2023). L'aire entre la courbe perturbée et la courbe de référence, l'amplitude maximale ou la durée de la perturbation peuvent être utilisées pour caractériser la réponse des animaux face à une perturbation (Taghipoor *et al.*, 2023). Cependant, ces méthodes sont applicables à des séries temporelles et nécessitent donc une collecte de données individuelles en continu à un pas de temps court et régulier afin de permettre la détection de changements dans la trajectoire des variables étudiées.

Le développement de l'élevage de précision offre la possibilité d'obtenir ces données et ainsi de faciliter l'évaluation du bien-être des animaux de rente en diminuant le temps nécessaire aux observations. Ce concept consiste en l'utilisation de technologies (*e.g.*, capteurs connectés, caméra, intelligence artificielle) afin d'aider l'éleveur dans le suivi de ses animaux et sa prise de

décision dans la conduite de l'élevage (Veissier *et al.*, 2019). Initialement développée dans l'objectif d'optimiser la rentabilité des exploitations et de faciliter le travail des éleveurs, l'acquisition de données par les capteurs ou les automates pourrait aussi être pertinente pour évaluer le bien-être des animaux tout au long de leur vie. Ces technologies peuvent notamment collecter en continu des données environnementales (*e.g.*, température, humidité), physiologiques (*e.g.*, température de l'animal, pH ruminal), zootechniques (*e.g.*, production laitière, poids, ingestion) et comportementales (*e.g.*, activité, interactions sociales). Un certain nombre de ces variables peuvent permettre une détection précoce de pathologies et ainsi guider les actions de l'éleveur afin d'améliorer la prise en charge des animaux. Par exemple, une boiterie peut s'accompagner d'une diminution du temps d'ingestion (Barker *et al.*, 2018). La diminution de la quantité de matière sèche ingérée et du temps de rumination, associée à une hausse de la température ambiante, peut indiquer un état de stress thermique aboutissant à une baisse de la production laitière (Corazzin *et al.*, 2021) ou encore la diminution de l'ingestion et du temps d'alimentation journalier peut permettre de détecter une cétose 3 à 5 jours avant l'apparition des symptômes identifiables par l'éleveur (González *et al.*, 2008). De la même façon, l'ingestion de matière sèche et la consommation d'eau sont aussi influencées par la période d'œstrus des vaches laitières et pourraient constituer des indicateurs pertinents pour identifier le moment le plus favorable à l'insémination des animaux (Reith *et al.*, 2014). Le temps que la vache passe debout ou couché peut également représenter un bon indicateur de l'état de bien-être et de santé de l'animal (Brochers *et al.*, 2016). Le temps couché représente une grande partie du budget-temps des vaches (*i.e.*, en moyenne 11 heures par jour, Gomez *et al.*, 2010) et il est influencé par de nombreux facteurs, comme le rang social, la qualité des zones de couchage (espace, litière) ou encore la présence de boiteries (Mattachini *et al.*, 2013).

Au-delà de l'identification ciblée de certaines pathologies, les données produites par ces technologies peuvent trouver un intérêt dans une évaluation plus globale du bien-être des animaux d'élevage. Durand *et al.* (2023) ont par exemple utilisé des données comportementales analysées à partir de vidéos et de distributeurs automatiques d'aliments (activité, comportements sociaux et comportements alimentaires) pour classer des truies gestantes selon leur niveau de bien-être avec une précision de 80%, grâce à des techniques d'analyse automatisées de machine learning. Cette approche semble prometteuse pour faciliter l'évaluation et le suivi en continu du bien-être animal en élevage (Durand *et al.*, 2023). Il existe peu d'études similaires chez les vaches laitières. Dans leurs travaux, Stygar *et al.* (2023) s'intéressent à l'évaluation du bien-être de vaches laitières en utilisant un algorithme d'apprentissage automatique ayant pour entrées des

données collectées par des accéléromètres et des compteurs à lait permettant de classer les animaux selon 3 niveaux de bien-être (faible, moyen, élevé) à partir de leur budget-temps (temps passé couché, en rumination, en alimentation, en déplacement) et de leur production laitière. Cette méthode s'est montrée plus performante pour classer les individus se trouvant dans un état de bien-être faible ou élevé que modéré révélant un manque de précision sur la détection d'altérations légères du bien-être des animaux ou d'erreurs dans la collecte des données. Selon les auteurs, l'utilisation d'un tel algorithme serait pertinent en complément d'une évaluation humaine du bien-être afin d'avoir un suivi en continu des animaux mais la précision du modèle seul n'est pas suffisante pour apporter à l'éleveur une aide à la décision individualisée sur ses animaux (Stygar *et al.*, 2023).

Malgré le développement et la pertinence des technologies utilisables en élevage pour faciliter la collecte des données et l'implication favorable qu'elles peuvent avoir pour la détection précoce de pathologies et l'évaluation du bien-être animal, analyser la réponse d'un animal en temps réel s'avère difficile. Cela demande de développer des méthodes fiables, précises et automatisées applicables en fermes commerciales permettant de détecter et prédire d'éventuelles pathologies ou difficultés d'adaptation des animaux (Ben Abdelkrim *et al.*, 2021 ; Nguyen-Ba *et al.*, 2019). De nombreuses études concernant l'alimentation, la reproduction ou encore le suivi de la santé des animaux en élevage laitier s'intéressent à l'utilisation de l'apprentissage automatique pour analyser en temps réel les données collectées automatiquement par les différents capteurs. Mais des recherches sont encore nécessaires pour améliorer les performances des algorithmes pour automatiser les procédures et faciliter la mise en œuvre de ces méthodes sur le terrain, notamment en intégrant plusieurs variables dans le processus afin de réduire la détection de faux positifs (Stygar *et al.*, 2023 ; Cockburn, 2020). D'autre part, les méthodes de lissage habituellement utilisées pour limiter le bruit et la variabilité des données pourraient être pertinentes pour détecter et caractériser les perturbations dans la trajectoire des variables d'intérêt à l'échelle de l'individu (Codrea *et al.*, 2011).

L'étude présentée dans ce rapport vise à déterminer si l'utilisation de données individuelles d'activité, de production laitière ou d'ingestion enregistrées en élevage combinées à des méthodes de détection de ruptures permettent d'identifier des perturbations liées à des événements observés par les animaliers (*e.g.*, ajustements de la ration, pathologies, chaleurs). Pour répondre à cet objectif, nous nous sommes focalisés sur l'utilisation de courbes de lissage en utilisant deux méthodes. La première a consisté en une approche par des B-splines cubiques pénalisés au travers du package R FDA pour l'analyse de données fonctionnelles (Ramsay *et al.*,

2024) et à l'étude d'indicateurs spécifiques sur la réponse des animaux face à la perturbation (*i.e.*, l'aire entre les courbes de lissage). La seconde a consisté en une approche par des lissages à noyaux au travers du package R UpDown (Rohmer *et al.*, 2023) qui se focalise sur l'information du groupe pour déterminer si un individu a fait face à une perturbation ou non. Les indicateurs fournis par le package UpDown reposent alors sur l'analyse des dérivées première de la courbe de lissage (*i.e.*, décélération, minimum locaux) pour quantifier le début, la fin et l'intensité des perturbations. Ces méthodes sont appliquées aux données de production laitière déjà couramment collectées en élevage laitier commercial, aux données d'ingestion journalière ainsi qu'au temps passé debout au vue de leur sensibilité face aux perturbations, notamment pathologiques. Les données sont obtenues à partir d'une expérimentation durant laquelle des vaches laitières ont été soumises à des ajustements hebdomadaires de la quantité d'aliment concentré distribué. La performance des méthodes est testée en réalisant des calculs de sensibilité et de spécificité pour chacune des variables considérées. Il était attendu que les méthodes utilisées devraient permettre de détecter les ruptures les plus importantes dans les trajectoires des données comportementales et zootechniques (rendement laitier et quantité de matière sèche ingérée) et que ces changements pourraient être reliés à des événements associés à la reproduction, à la santé ou à l'alimentation observés par les animaliers.

Matériel et méthodes

1. Note éthique

L'expérimentation qui a fourni les données utilisées dans ce projet s'est déroulée entre Septembre 2021 et Avril 2022 à la ferme expérimentale de l'Institut National de la Recherche pour l'Agriculture, l'Alimentation et l'Environnement (INRAE) à Méjusseume (Le Rheu, Bretagne, France, <https://doi.org/10.15454/yk9q-pf68>, no. C-35–275-23). Les procédures ont été effectuées en accord avec la Directive 2010/63/UE du Parlement européen et du Conseil de l'Union européenne, relative à la protection des animaux utilisés à des fins scientifiques et approuvées en Septembre 2021 par un comité d'éthique local ainsi que par la Ministère de l'Enseignement Supérieur et de la Recherche (référence APAFIS #31836-2021053017181790 v3).

2. Animaux et hébergement

L'expérimentation a été menée sur 40 vaches laitières de race Prim'holstein durant leur deuxième (N = 23) ou troisième lactation (N = 17). Les animaux sont entrés dans la procédure expérimentale entre Septembre et Novembre 2021, le jour de leur vêlage ou le lundi suivant si celui-ci se déroulait durant le week-end. Les animaux sont restés dans l'expérimentation pendant environ 4 mois à partir de leur date d'entrée dans la procédure, avec une moyenne de 125 jours par individu (± 2 jours).

Les vaches étaient hébergées en groupes et dans un bâtiment pendant toute la durée de l'expérience. Elles avaient un accès *ad libitum* à l'eau. La ration était distribuée dans des mangeoires individualisées limitant l'accès à une vache spécifique grâce à des portillons (Proval, Beaupréau-en-Mauges, France) et alimentées *via* un distributeur dit roulimètre capable de distribuer à chaque vache une ration calculée et programmée individuellement (Prototype INRAE). Les vaches étaient équipées d'un collier contenant une puce RFID (BouMatic, Saint Rémy-lès-Chevreuse, France) permettant une reconnaissance individuelle aux mangeoires et d'un accéléromètre (Medria, Saint-Berthevin, France) pour mesurer leur activité physique. Les animaux étaient nourris *ad libitum* avec une ration de base dite semie-complète constituée d'un mélange de fourrages et de concentré. Cette ration était distribuée deux fois par jour en quantité équivalente après la traite. Les vaches étaient traitées à 7:00 et 16:00 *via* un système de traite rotative automatique (Delaval France, Elancourt, France).

3. Procédure expérimentale

Les animaux ont été répartis dans deux groupes expérimentaux (AS : alimentation standard et AM : alimentation sur mesure). Selon leur traitement expérimental, les vaches recevaient un apport supplémentaire de concentré (fabriqué par VALOREX) et réparti sur les deux distributions journalières. Les vaches du groupe AS étaient nourries avec la ration de base *ad libitum* complétée d'un apport fixe de 3 kg de concentré par jour. Le groupe AM était nourri avec la ration de base *ad libitum* complétée d'un apport de concentré variable allant de 0 à 6 kg par jour en fonction du gain de poids de l'animal durant la semaine précédente (Tableau 1).

La première semaine de lactation correspondant à la phase colostrale a duré entre 4 et 11 jours après le vêlage (en moyenne 8 jours). Durant cette période, toutes les vaches ont été nourries avec la ration de base et un apport supplémentaire de 3 kg de concentré d'ajustement par jour.

Les animaux ont ensuite été répartis dans les deux groupes expérimentaux (AS et AM) selon leur parité, la date du vêlage et leur gain de poids au cours de la première semaine de lactation, à raison de 20 vaches par groupe. L'ajustement de l'apport de concentrés du groupe AM se faisait tous les mardis dès la deuxième semaine de lactation. La différence d'apport de concentré supplémentaire ne pouvait excéder 3 kg d'une semaine à l'autre afin d'éviter des changements trop rapides.

Tableau 1. Règle d'ajustement des concentrés selon le gain de poids individuel lors de la semaine précédente. La variation d'apport de concentrés entre 2 semaines ne peut excéder 3 kg, même si le gain de poids nécessite un ajustement plus important.

| | Poids de la semaine précédente | Quantité de concentré d'ajustement distribuée |
|----------------------|---------------------------------|---|
| Si perte de poids | ≥ -15 kg | 6 kg |
| | $-8 \text{ kg} \leq x < -15$ kg | 5 kg |
| | $-1 \text{ kg} \leq x < -8$ kg | 4 kg |
| Gain de poids stable | $+1 \text{ kg} \leq x < -1$ kg | 3 kg |
| Si gain de poids | $+1 \text{ kg} \leq x < +8$ kg | 2 kg |
| | $+8 \text{ kg} \leq x < +15$ kg | 1 kg |
| | $\geq +15$ kg | 0 kg |

4. Mesures zootechniques et comportementales

La production laitière journalière individuelle (PL) a été enregistrée automatiquement *via* la machine à traire deux fois par jour. La quantité de matière sèche ingérée par jour (MSI) de chaque vache a été déterminée à partir du poids des refus, la proportion fourrage/concentré distribué ainsi que des teneurs en matière sèche des aliments.

Des données comportementales ont également été collectées *via* des accéléromètres fixés sur un collier autour du cou de l'animal. Ces appareils ont enregistré en continu le comportement de l'animal et déterminé l'activité principale effectuée sur des périodes de 5 minutes pendant l'intégralité de l'expérience. Des données interprétées ont été fournies par l'entreprise MEDRIA permettant d'identifier le temps d'ingestion, de rumination, de repos (debout ou couché sans distinction), de suractivité, le temps passé dans d'autres activités ainsi que le temps passé debout (DEB) pour chaque individu.

Le temps passé debout semble être fiablement mesuré par différents types d'accéléromètres (Brochers *et al.*, 2016). La suite de l'analyse portera donc uniquement sur ce comportement.

Les évènements sanitaires et les chaleurs observés par le personnel de l'élevage ainsi que les inséminations et les ajustements alimentaires identifiés à partir du protocole d'expérimentation décrit ci-dessus ont été consignés pendant l'expérimentation. Ils constituent une liste d'évènements identifiés par les méthodes de référence (nommé gold standard ci-après).

5. Analyses statistiques

5.1 Nettoyage des données et analyse descriptive

Les données ont été analysées avec le logiciel R version 4.3.2 (R Core Team, 2023) et nettoyées *via* le package Dplyr (Wickham *et al.*, 2023). Ainsi, les données collectées en dehors de la période d'expérimentation ont été écartées de l'analyse et les données considérées comme aberrantes ont été supprimées. Deux jeux de données ont donc été constitués, d'une part les données zootechniques (PL et MSI), et d'autre part les données comportementales (DEB).

La PL et la MSI utilisées dans les analyses ont été agrégées par individu et par jour en sommant les différents relevés enregistrés dans la journée. Afin de ne pas avoir d'interruption dans la trajectoire des données, les périodes de données manquantes ne dépassant pas 3 jours successifs ont été interpolées *via* un lissage de Loess (Cleveland *et al.*, 1992). Ces données manquantes représentaient 185 jours pour la PL et 209 jours pour la MSI sur l'ensemble des vaches.

Concernant les données comportementales issues des accéléromètres, un tableau agrégé a été construit à partir des données fournies par MEDRIA pour aboutir à un budget-temps individuel exprimé en minutes par jour (ruminantion, ingestion, repos, suractivité, autres activités, position debout). Les journées dont le temps d'activité totale n'égalait pas 1440 minutes ainsi que les périodes de relevées inférieures à 15 jours successifs ont été considérées comme manquantes (N = 384 jours). Deux individus ayant une longue période de données manquantes ne peuvent être interpolées (N = 7 jours et N = 17 jours) ont été exclus des analyses pour la variable comportementale uniquement.

Une variable "jours en lactation" a été calculée en faisant la différence entre la date de l'observation et la date de vêlage de chaque individu.

L'effet de la stratégie d'alimentation sur les variables est analysé *via* un modèle linéaire mixte *via* le package lme4 (Bates *et al.*, 2024). Les conditions d'application ont été validées. Les individus ont été définis comme un effet aléatoire, le groupe expérimental (AM ou AS) comme effet fixe.

5.2 Méthodes de lissage : FDA et UpDown

Les méthodes de lissage permettent de décrire la dynamique de données continues en distinguant les variations issues de perturbations de celles issues de la variabilité des données (*e.g.*, variations interindividuelles ou journalières, erreur de mesure). Souvent utilisés pour réduire le bruit et faciliter l'analyse des effets de différents traitements, les lissages peuvent être utilisés pour identifier des perturbations dans la trajectoire des données et les caractériser (Codrea *et al.*, 2011).

Pour des données complexes et non périodiques, les B-splines cubiques sont couramment utilisées comme fonctions de bases et peuvent être définies au travers du package R FDA. Ces fonctions de bases précisément décrites dans Codrea *et al.* (2011), sont définies au travers d'une succession de courbes polynomiales (ici de degré 3), estimées à partir des données observées. L'ajustement de la B-spline aux données observées s'effectue au travers de l'ajout d'une pénalité sur la dérivée seconde du spline, spécifiée par un paramètre de lissage λ . Lorsque λ est faible, la B-spline s'ajuste finement aux données observées et permet la détection de perturbations à court terme. Lorsqu'il est élevé, le poids de la pénalité appliquée à la dérive augmente, le lissage devient plus important permettant ainsi la détection de perturbations à long terme (Ben Abdelkrim *et al.*, 2021 ; Codrea *et al.*, 2011 ; Levitin *et al.*, 2007).

Dans la présente étude, le lissage a été appliqué à des données corrigées par les jours de lactation afin de tenir compte de l'évolution naturelle des variables durant la lactation. Pour ce faire, une valeur d'observation médiane a été calculée sur l'ensemble du troupeau pour chaque jour de lactation, puis soustraite aux données individuelles de ce même jour. Plusieurs niveaux de lissage ont été testés en faisant varier la valeur du paramètre λ à 2,5, 10^4 et 10^6 . Afin de pouvoir comparer les individus entre eux, ce même paramètre a été appliqué aux individus pour les 3 variables. L'aire entre les courbes (*i.e.*, pour une même perturbation, la somme des différences entre les deux courbes de lissage divisée par la durée de la perturbation) a été utilisée comme indicateur de l'importance de la perturbation. Ainsi, chaque différence entre les deux courbes a conduit à l'identification d'un candidat à une potentielle perturbation. Afin de sélectionner les candidats les plus susceptibles de correspondre à une véritable perturbation, deux seuils ont été appliqués (à

10% et 2,5%) pour sélectionner les valeurs absolues des aires les plus fortes. Les perturbations candidates incluses dans cette sélection sont considérées comme les perturbations détectées par FDA pour la suite de l'analyse. Le seuil de 10% a été choisi arbitrairement, celui à 2,5% a été utilisé pour permettre de sélectionner un nombre de perturbations similaires à ce qui a été détecté par l'autre méthode.

La seconde méthode utilisée dans cette étude a été la méthode UpDown, utilisant le principe d'un lissage à noyau. Elle permet la détection de perturbations dans la trajectoire de données longitudinales organisées en différents niveaux hiérarchiques (*e.g.*, individus, groupe, troupeau). Dans un premier temps, l'algorithme construit une trajectoire médiane (*i.e.*, sans les perturbations potentielles) à partir des données observées *via* une courbe de lissage de Nadaraya-Watson. Un modèle de mélange Gaussien est ensuite utilisé pour étudier la distribution des valeurs minimales des dérivées premières de la courbe de lissage correspondant à la pente minimale de la trajectoire des données. Cette méthode de clustering non supervisée permet d'identifier les individus ayant subi une perturbation de ceux n'en ayant pas subi. Dans le cadre de l'élevage laitier par exemple, cette approche permet de distinguer une perturbation présente uniquement à l'échelle individuelle d'une autre provenant d'un événement ayant eu un impact visible sur un ensemble d'individus, voire sur le troupeau entier. Dans un second temps, l'algorithme valide les éléments détectés à un niveau hiérarchique supérieur en comparant le début de la perturbation pour chaque individu concerné. N'ayant qu'un niveau hiérarchique dans cette étude (*i.e.*, l'individu), cette seconde étape n'a pas été effectuée. L'importance du lissage (*h.int*) est précisé par la largeur de bande basée sur la longueur des observations au niveau hiérarchique le plus bas. Cette méthode permet d'obtenir le moment de début et de fin de la perturbation ainsi que son intensité (Le *et al.*, 2024 ; David *et al.*, 2023).

Dans la présente étude, plusieurs niveaux de lissage ont été testés en faisant varier le paramètre *h.int* à 5 et 3. Le choix du niveau de lissage a été fait par observation graphique. Une correction par les jours en lactation a été appliquée aux trois variables de la même façon que pour la méthode des B-splines.

L'utilisation des deux méthodes a abouti à l'obtention d'une liste de perturbations pour chacune des variables, automatiquement pour UpDown (classification non-supervisée) ou définit selon le seuil déterminé pour FDA (classification supervisée). Avec UpDown, des sorties graphiques ont également été obtenues permettant de visualiser les données individuelles des vaches, les courbes de lissage ainsi que les perturbations détectées. Des figures similaires ont été codées sur R en utilisant la méthode par B-splines (Figure 1).

Les perturbations détectées par UpDown et FDA avec le seuil à 2,5% ont été manuellement comparées entre elles afin de vérifier si les deux méthodes détectaient les mêmes perturbations. Pour cela, un décalage de 2 jours avant ou après le début de la perturbation détectée par FDA était toléré pour considérer qu'elle correspondait à celle détectée par UpDown (Le *et al.*, 2024). Ces perturbations ont ensuite été comparées au gold standard afin de vérifier si celles-ci étaient liées ou non à un événement observé par les animaliers.

5.3 Calcul de sensibilité et spécificité

Les perturbations détectées par les méthodes de lissage ont été comparées aux événements sanitaires (N = 239), alimentaires (N = 112) ou de reproduction (N = 78) identifiés par les méthodes de références (*e.g.*, observations par les animaliers, diagnostics vétérinaires). Le gold standard ainsi constitué contenait 429 événements. Cette comparaison a été faite sur l'ensemble du gold standard, sans distinction du type d'événement.

La date de début d'une perturbation détectée pouvant ne pas être exactement la même que celle de l'événement observé par les animaliers, une règle pour associer une perturbation à un événement du gold standard a été établie. Ainsi, une perturbation était considérée comme liée à un événement du gold standard lorsque la date de début de la perturbation se situait dans un intervalle autour de la date d'observation faite par l'animalier. Pour que cet intervalle conserve un sens biologique, il a été adapté à chaque type d'événement : 4 jours avant et après un événement sanitaire, 1 jour avant et 4 jours après une insémination afin de prendre en compte l'oestrus et de 4 jours uniquement après un changement de ration (Figure 2).

Lorsque la perturbation détectée par l'une des méthodes correspondait à un événement du gold standard, les jours correspondant ont été classés comme vrais positifs (VP). S'il n'y avait pas d'événement enregistré sur la même période que la perturbation, les jours correspondant à celle-ci ont été classés comme faux positifs (FP). Lorsque la méthode ne détectait pas de perturbation mais qu'un événement est enregistré dans le gold standard, les jours correspondant ont été classés comme faux négatifs (FN). Lorsqu'aucun événement n'a été enregistré et qu'aucune perturbation n'a été détectée, les jours correspondant ont été classés comme vrais négatifs (VN). A partir de ces paramètres, la sensibilité (*i.e.*, capacité de la méthode à détecter des perturbations : $VP/(VP + FN)$) et la spécificité (*i.e.*, capacité de la méthode à ne pas détecter à tort des perturbations : $VN/(VN + FP)$) des méthodes ont été calculées pour chaque variable.

Plus la sensibilité est élevée, plus la méthode est capable de détecter les vrais positifs (*i.e.*, les perturbations détectées sont bien associées aux événements du gold standard). D'autre part, plus la spécificité est élevée, plus la méthode évite la détection de faux positifs (*i.e.*, les perturbations détectées ne correspondent à aucuns événements du gold standard, Chu, 1999).

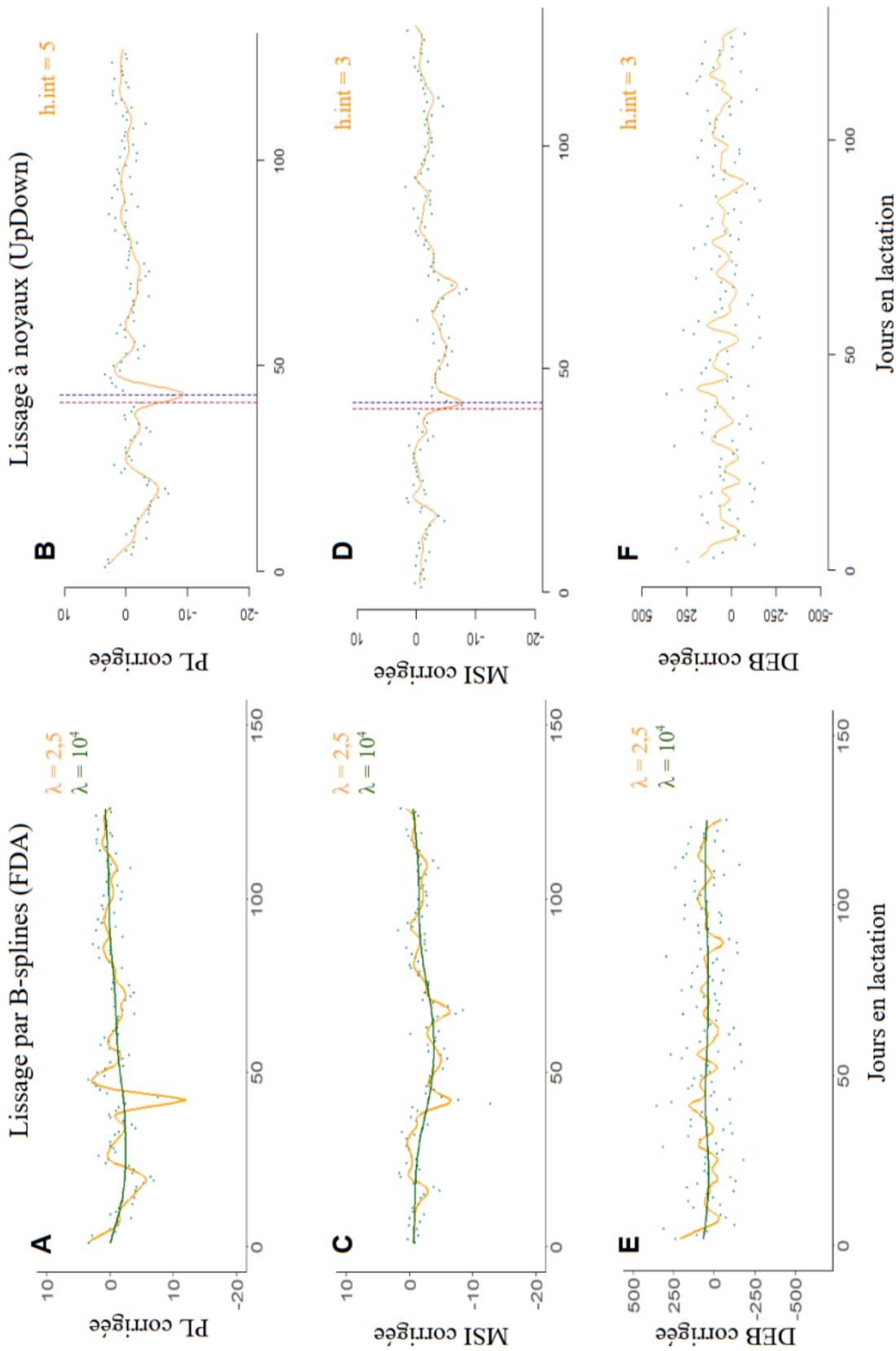
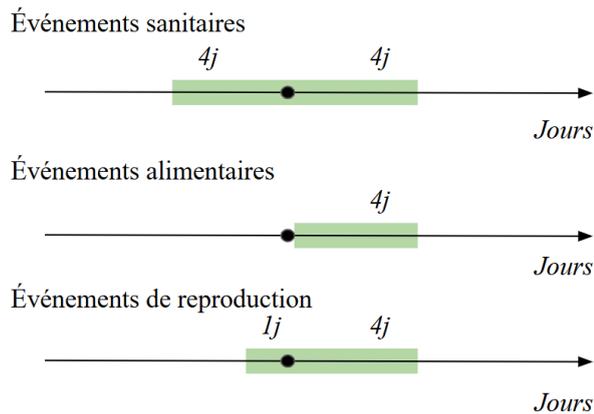


Figure 1. Graphique des lissages effectués par les méthodes FDA et UpDown pour la production laitière (PL, respectivement A et B), pour la quantité de matière sèche ingérée (MSI, respectivement C et D) et le temps passé debout (DEB, respectivement E et F) pour l'une des vaches du troupeau. Les variables sont corrigées par les jours de lactation. Les verticales en pointillés correspondent au début (en rouge) et à la fin (en bleu) de la perturbation détectée par UpDown. Les paramètres λ (FDA) et h_int (UpDown) permettent de définir le niveau de lissage des courbes.

Règle d'association événements-perturbations



● Événement enregistré dans le gold standard

Correspondance entre le gold standard et les méthodes

| Gold standard | Détection par une méthode | |
|---------------|---------------------------|------|
| Événement | Perturbation | → VP |
| NA | Perturbation | → FP |
| NA | NA | → VN |
| Événement | NA | → FN |

Figure 2. Règle d'association entre un événement du gold standard et une perturbation détectée par les méthodes (B-splines, UpDown) et de la détermination des vrais positifs (VP), faux positifs (FP), vrais négatifs (VN) et faux négatifs (FN). La règle d'association est définie à 4 jours autour de l'événement observé par l'animalier, excepté pour les événements liés à la reproduction qui se limite à 1 jour avant l'événement ce qui correspond à la période d'oestrus de la vache.

Résultats

1. Effet de la stratégie d'alimentation sur la production laitière, l'ingestion et le temps passé debout

L'analyse s'est effectuée sur un jeu de données contenant 4979 observations pour la PL, 4985 pour la MSI et 4383 pour DEB (Tableau 3).

Le modèle linéaire mixte n'a pas mis en évidence de différence significative entre les deux groupes expérimentaux (AM et AS) pour aucune des trois variables (PL : LMM type 2, $p = 0,3935$; MSI : LMM type 2, $p = 0,87$; DEB : LMM type 2, $p = 0,8947$). L'effet de la stratégie d'alimentation n'a donc pas été davantage analysée par la suite.

Tableau 3. Description des variables journalières étudiées : temps passé debout (DEB), quantité de matière sèche ingérée (MSI) et production laitière (PL) pour les deux groupes expérimentaux AM (alimentation sur mesure) et AS (alimentation standard). ES : erreur standard.

| Groupes expérimentaux | DEB | | MSI | | PL | |
|-----------------------|------------------------|------------------------|--------------------------|--------------------------|-------------------------|-------------------------|
| | AM | AS | AM | AS | AM | AS |
| Nombre d'individus | 18 | 20 | 20 | 20 | 20 | 20 |
| Nombre d'observations | 2109 | 2274 | 2505 | 2480 | 2501 | 2478 |
| Moyenne \pm ES | 902 \pm 116 min/j | 902 \pm 115 min/j | 24,55 \pm 3,50 kg/j | 24,48 \pm 3,68 kg/j | 38,00 \pm 5,08 L/j | 37,99 \pm 5,38 L/j |

2. Choix du niveau de lissage

Pour FDA, trois niveaux de lissage ont été testés : $\lambda = 2,5$, $\lambda = 10^4$ et $\lambda = 10^6$.

Le lissage était plus important pour $\lambda = 10^4$ et $\lambda = 10^6$ que pour $\lambda = 2,5$ ce qui a permis de définir une courbe de référence avec les lissages les plus forts et de conserver une courbe perturbée avec le lissage le plus faible (Figure 3). Afin de tenir compte de l'évolution normale des variables au cours de la lactation, le lissage le plus important (*i.e.*, $\lambda = 10^6$) n'a pas été retenu pour les analyses. Par observation graphique appuyée par une précédente étude (Codrea *et al.*, 2011), le choix des niveaux de lissage pour la suite des analyses s'est donc fait sur les valeurs de $\lambda = 2,5$ et $\lambda = 10^4$. Le résultat était similaire pour la PL, la MSI et le temps passé debout. Le même choix de lissage a donc été fait pour les trois variables et appliqué de la même façon sur tous les individus.

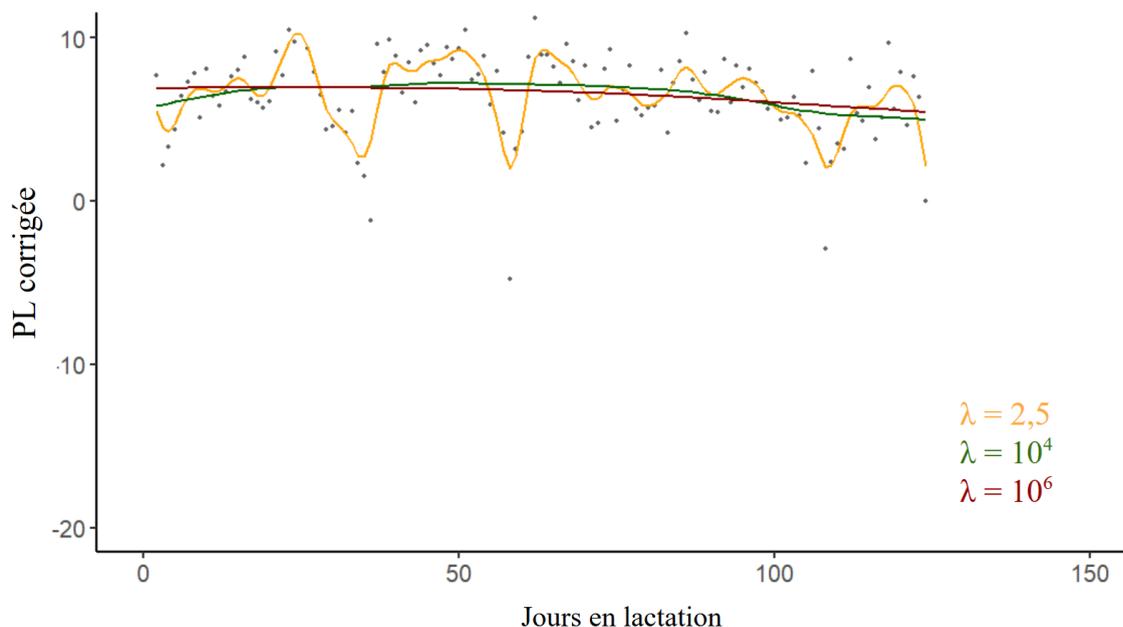


Figure 3. Lissages par B-splines (FDA) appliqués à la production laitière (PL) corrigée par les jours en lactation pour l'une des vaches du troupeau. Trois niveaux de lissage sont testés en faisant varier le paramètre λ à 2,5, 10^4 et 10^6 . Seuls les lissages obtenus avec $\lambda = 2,5$ et $\lambda = 10^4$ sont retenus pour les analyses.

Pour UpDown, deux niveaux de lissage ont été testé : h.int = 5 et h.int = 3.

Le lissage obtenu avec le paramètre h.int = 3 était moins important que pour h.int = 5 ce qui donnait une courbe de lissage qui suivait davantage les variations journalières des données pour le paramètre le plus faible. De plus, le niveau de lissage entraînait une différence de détection de perturbations. Le lissage le plus important (h.int = 5) permettait de détecter 17 perturbations pour la PL, 8 pour la MSI et 3 pour DEB. Le lissage plus faible (h.int = 3) permettait quant à lui la détection de 2 perturbations pour la PL, 10 pour la MSI et 3 pour DEB. Ces résultats mettent en avant l'importance d'ajuster les paramètres de lissage à la variable étudiée. Définir le paramètre h.int à 5 pour la PL et 3 pour la MSI et DEB semblait permettre de concilier la réduction de la variabilité des données entre les jours d'observation et la mise en évidence de chutes dans la trajectoire des données (Figure 4). La suite de l'analyse a donc été faite avec ce niveau de lissage distinct entre les variables, mais identique pour tous les individus.

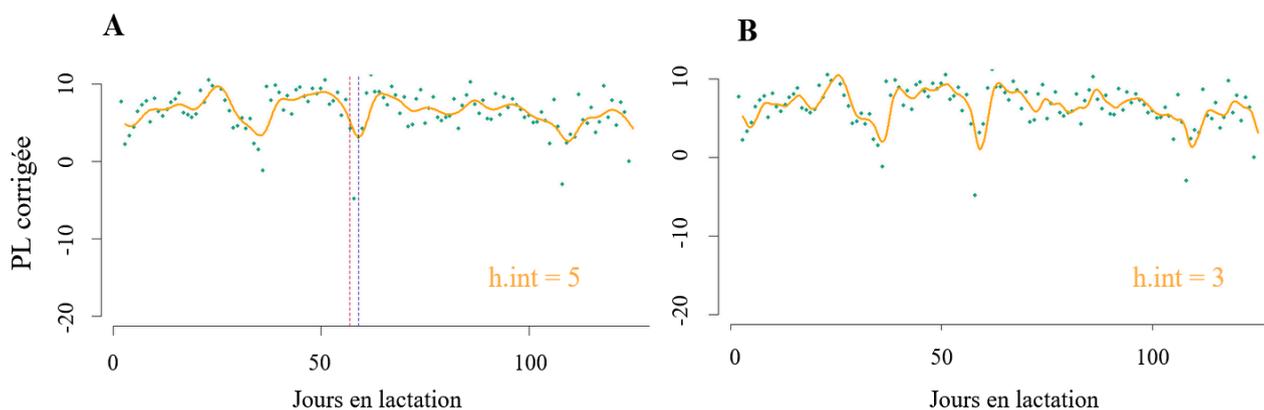


Figure 4. Lissages effectués par la méthode UpDown appliqués à la production laitière (PL) corrigée par les jours en lactation pour l'une des vaches du troupeau. Deux niveaux de lissage sont testés en faisant varier le paramètre $h.int$ à 5 (**A**) et 3 (**B**). Les verticales en pointillées présentent sur le graphique A correspondent au début (en rouge) et la fin (en bleue) d'une perturbation détectée par UpDown sur cet individu.

3. Détection des perturbations

Sans définir de seuil, FDA détectait plus de perturbations que UpDown. En fixant un seuil à 2,5%, le nombre de perturbations détectées par FDA est assez proche de UpDown, exceptée pour la variable comportementale (DEB). En effet, UpDown ne détectait que 3 perturbations quand FDA en détectait 20 (Tableau 4).

Tableau 4. Nombre de perturbations détectées par les deux méthodes de lissage (B-splines et UpDown) appliquées au temps passé debout (DEB), à la quantité de matière sèche ingérée (MSI) et à la production laitière (PL).

| | DEB | MSI | PL |
|--|-----|-----|-----|
| Lissage par B-spline (FDA) | 824 | 714 | 723 |
| Lissage par B-spline (FDA) avec un seuil à 10% | 82 | 71 | 72 |
| Lissage par B-spline (FDA) avec un seuil à 2,5% | 20 | 19 | 18 |
| Lissage à noyaux (UpDown) | 3 | 10 | 17 |

Pour UpDown, aucune des perturbations détectées par la méthode dans la variable DEB ne correspondait à un événement du gold standard. Pour la MSI, 3 perturbations correspondaient au gold standard sur les 10 détectées par la méthode (soit 30% de vrais positifs). Pour les 17 perturbations détectées sur la PL, 5 d'entre elles correspondaient à un événement du gold standard (soit 30% de vrais positifs). Pour FDA avec le seuil à 2,5%, sur les 20 perturbations détectées par la méthode sur la variable comportementale, 13 correspondaient à un événement du gold standard (soit 65% de vrais positifs). Pour la MSI, il y avait également 13 correspondances sur les 19 perturbations (soit 68% de vrais positifs) et 10 correspondances sur les 18 perturbations détectées sur la PL (soit 55% de vrais positifs).

Sur les 86 perturbations détectées par UpDown et FDA (avec le seuil à 2,5%), les perturbations n'étaient pas les mêmes entre les deux méthodes. Exceptée pour 2 perturbations sur 2 vaches différentes. L'une des perturbations était détectée par les deux méthodes sur la PL, l'autre était détectée à la fois sur la PL par UpDown et sur la variable comportementale par FDA.

4. Calculs de sensibilité et spécificité des méthodes

Quelle que soit la variable considérée, FDA se montre plus sensible que UpDown pour détecter les jours perturbés (Figure 5a). Inversement, la spécificité est plus importante pour UpDown que pour FDA, quel que soit le seuil choisi, et ce pour les 3 variables. Appliquer un seuil à 10% pour FDA rend la méthode plus sensible qu'à 2,5% mais de fait moins spécifique.

Il n'y a pas de différence marquée de sensibilité ou de spécificité entre les variables zootechniques (production laitière et quantité de matière sèche ingérée). En revanche, le temps passé debout montre une sensibilité plus faible que les variables zootechniques avec FDA pour les deux seuils.

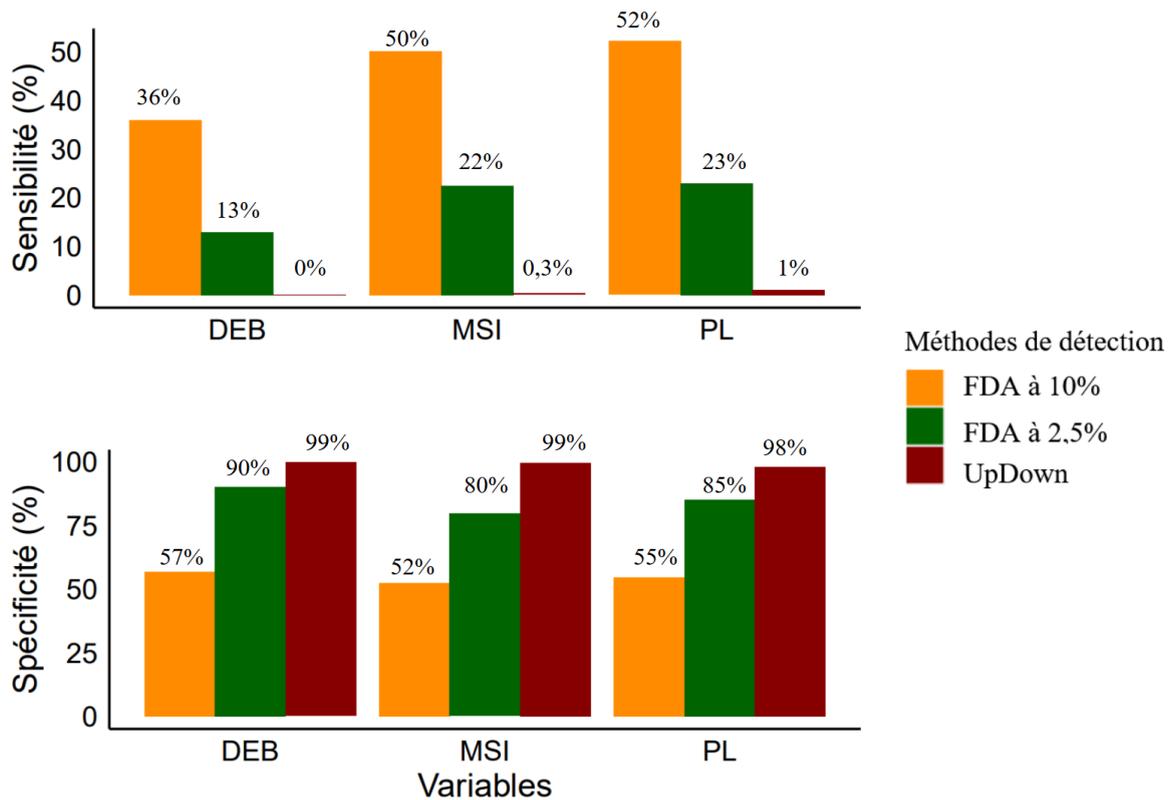


Figure 5a. Spécificité et sensibilité des méthodes UpDown et FDA (seuil à 10% et 2,5% sélectionnant les perturbations les plus importantes) pour les 3 variables : production laitière (PL), la quantité de matière sèche ingérée (MSI) et le temps passé debout (DEB).

[paragraphe ajouté après soumission du rapport]

Les calculs de sensibilité et spécificité ont été refaits en perfectionnant la méthode d'identification des vrais positifs sur R. Ainsi, des erreurs liées au décalage entre les dates de début et de fin des perturbations par rapport aux événements du gold standard ont pu être corrigées. De fait, davantage de jours considérés comme vrais positifs sont identifiés et par conséquent, cela entraîne moins de faux négatifs et de faux positifs.

Finalement, les performances de UpDown n'ont pas nettement été modifiées par cette correction. La sensibilité plus importante de FDA comparée à UpDown se confirme bien que celle-ci ait nettement diminué pour FDA (*e.g.*, de 23% à 6% pour la production laitière). En revanche, la spécificité de FDA a nettement augmentée pour les trois variables étudiées (*e.g.*, de 85% à 99,6% pour la production laitière), se rapprochant ainsi de celle de UpDown (Figure 5b).

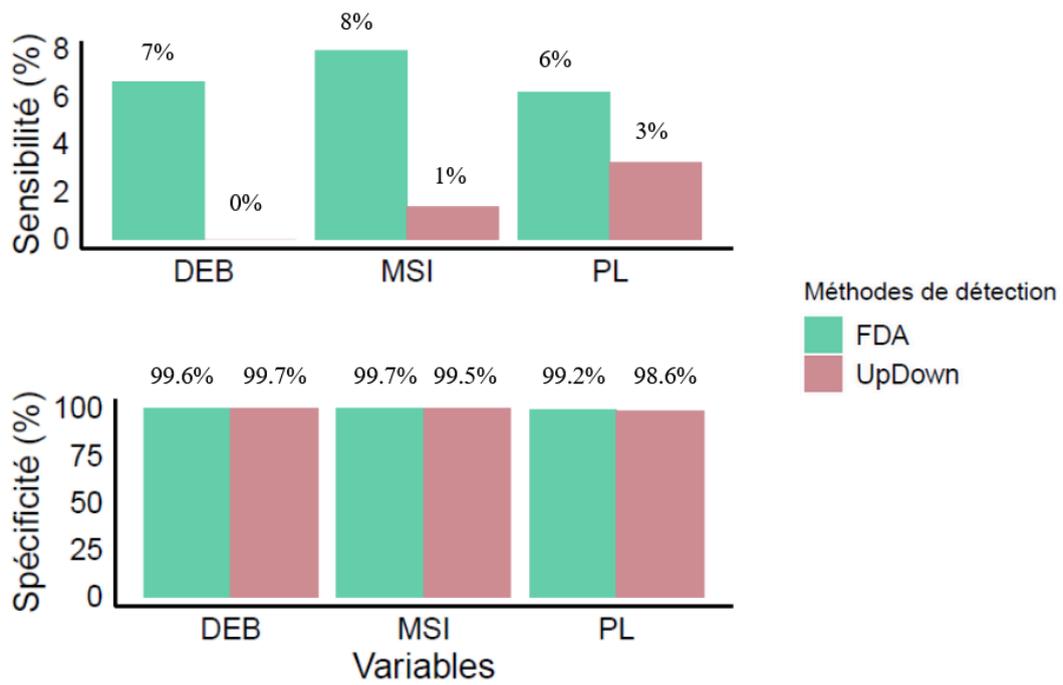


Figure 5b. Spécificité et sensibilité des méthodes UpDown et B-splines *via* le package FDA (uniquement avec le seuil à 2,5% sélectionnant les perturbations les plus importantes) pour les 3 variables : production laitière (PL), la quantité de matière sèche ingérée (MSI) et le temps passé debout (DEB).

Discussion

Les deux méthodes FDA et UpDown permettent la détection de changements dans la trajectoire des données de production laitière (PL), d'ingestion (MSI) et du temps passé debout (DEB). FDA détecte davantage de perturbations que UpDown. L'analyse des méthodes montre que FDA est plus sensible que UpDown, c'est-à-dire qu'il permet une meilleure détection des vrais positifs (*i.e.*, correspondance entre la perturbation détectée et un événement du gold standard). Mais la plus forte spécificité de ce dernier le rend plus performant pour éviter la détection de faux positifs (*i.e.*, détection d'une perturbation qui n'est pas liée à un événement enregistré dans le gold standard). De plus, UpDown ne détecte que 3 perturbations sur la variable DEB pour l'ensemble du troupeau, tandis que c'est pour cette variable que FDA détecte le plus de perturbations (N = 824 sans seuil défini et N = 20 avec un seuil à 2,5% des perturbations les plus importantes).

1. Fonctionnement, paramétrage et application des méthodes

Cette différence de performance entre les méthodes peut s'expliquer par la façon dont fonctionne les packages. En effet, avec FDA, l'aire entre les deux courbes de lissage est utilisée pour identifier les perturbations alors que UpDown se base sur l'intensité de la pente lors d'une chute de la trajectoire. De ce fait, FDA permet de détecter toutes les variations dans la trajectoire des données qu'elles soient décroissantes ou croissantes (*e.g.*, une diminution ou une augmentation de la production laitière) alors que UpDown ne détecte que les perturbations décroissantes, c'est-à-dire une chute dans la trajectoire de la variable (Le *et al.*, 2024). De plus, UpDown contient une approche statistique pour définir un seuil (*via* le modèle de mélange Gaussien) visant à différencier les variations relevant d'une perturbation de celles liées à la variabilité des données. Ainsi, cette méthode ne va détecter que les changements les plus importants tandis que FDA, n'intégrant pas la définition de ce seuil qui doit être fixé manuellement, détecte toutes les perturbations qu'elles soient mineures ou majeures.

Dans cette étude, le choix des seuils pour FDA a été fait de manière arbitraire, le seuil à 2,5% permettant de détecter un nombre de perturbations similaire à UpDown. Le choix du seuil pour FDA nécessite d'être exploré davantage car cela conditionne le nombre et l'importance des perturbations qui ont été utilisées pour les calculs de sensibilité et de spécificité. De la même façon, le choix des paramètres de lissage, en particulier l'importance de celui-ci, influence également le nombre de perturbations détectées par les méthodes et donc le calcul de leur performance (Le *et al.*, 2024 ; Ben Abdelkrim *et al.*, 2021).

La forte spécificité obtenue avec UpDown semble toutefois cohérente avec la littérature (Le *et al.*, 2024). Dans leur étude, Le *et al.* ont testé la performance de UpDown en utilisant des données de poids et d'ingestion chez le porc à différents niveaux hiérarchiques (individu, case, bande). Ils montrent une spécificité supérieure à 95% et une sensibilité qui augmente avec le niveau hiérarchique de 23% à 32% pour les perturbations individuelles et de 77% à 88% pour les perturbations des bandes (*i.e.*, niveau hiérarchique le plus élevé). Dans notre étude, un seul niveau hiérarchique est utilisé (*i.e.*, individu), malgré cela, la sensibilité obtenues pour les 3 variables avec UpDown est très inférieure (moins de 2%) aux résultats obtenus dans la précédente étude. Des difficultés ont été rencontrées pour automatiser le calcul de la sensibilité et de la spécificité sur R, notamment sur l'identification des vrais négatifs et des vrais positifs ce qui a pu fortement influencer les résultats de ces calculs. Malgré tout, la comparaison manuelle entre les perturbations détectées par les méthodes et les événements du gold standard semble indiquer que UpDown détecte moins de perturbations, mais aussi moins de vrais positifs que FDA ce qui tend à confirmer la faible sensibilité et la forte spécificité de UpDown.

Ces résultats peuvent s'expliquer par le fait que UpDown a été développée dans l'objectif de détecter des perturbations importantes à l'échelle du groupe. Si les variables utilisées dans notre étude, notamment la variable comportementale, ne sont pas fortement perturbées par un événement environnemental ou sanitaire, le package ne détecte pas ces perturbations mineures ce

qui peut conduire à l'identification de faux négatifs en comparaison avec le gold standard, et ainsi une faible sensibilité.

Il serait donc intéressant de recalculer la sensibilité et la spécificité de FDA et UpDown en testant davantage de paramètres, que ce soit le degré du lissage ou le seuil de caractérisation d'une perturbation, mais aussi en testant leurs performances sur d'autres variables, comme le poids des animaux par exemple (Ben Abdelkrim *et al.*, 2021).

2. Robustesse des animaux et choix des variables étudiées

Tenir compte de l'effet réel des événements potentiellement perturbants (*e.g.*, pathologies, changements alimentaires) auxquels sont exposés les animaux ainsi que de leur capacité à s'y adapter sans changements apparents dans leurs comportements ou leurs performances de production semble aussi nécessaire pour valider ces méthodes. En effet, des animaux robustes (*i.e.*, capable de minimiser l'impact d'une perturbation et de revenir rapidement à leur état initial, Nguyen-Ba *et al.*, 2020) présentent moins de changements dans la trajectoire des variables étudiées que des animaux moins robustes, ce qui peut impacter la sensibilité des méthodes.

Également, par observation des nuages de points, on peut se rendre compte de la variabilité journalière importante dans le temps que la vache passe debout. Cela peut être liée aux nombreux facteurs pouvant influencer cette variable (*e.g.*, le rang social de la vache, la configuration du bâtiment, les conditions environnementales (température, humidité), la conduite d'élevage (heure et fréquence des traites ou de la distribution des aliments), Mattachini *et al.*, 2017 ; Mattachini *et al.*, 2013). La sensibilité de FDA pour cette variable est inférieure à celle obtenue pour la production laitière ou l'ingestion ce qui traduit une faible association entre les perturbations détectées et les événements du gold standard. Bien que le temps que l'animal passe debout ou couché puisse être un indicateur de l'état de bien-être de l'animal (Brochers *et al.*, 2016), celui-ci ne semble pas permettre d'identifier un événement ponctuel en particulier en

utilisant ces méthodes et s'avère peu prédictif d'une perturbation liées à un événement enregistré dans le gold standard.

En revanche, l'ingestion et la production laitière semblent davantage pouvoir être liées à une perturbation, notamment sanitaire ou environnementale, comme l'ont déjà montré plusieurs études (Corazzin *et al.*, 2021 ; Green *et al.*, 2002). Cependant, la détection d'une perturbation sur la quantité de lait produit par exemple, se fait souvent à posteriori. Il serait donc pertinent d'effectuer davantage de recherches sur la détection en temps réel de ces perturbations pour anticiper une altération importante du bien-être de l'animal et/ou une baisse de production afin de mettre en place les mesures nécessaires (*e.g.*, traitement d'une pathologie, González *et al.*, 2008).

3. Fiabilité des données collectées par les capteurs et les automates

Une autre part importante des performances de ces méthodes relève de la fiabilité et de la pertinence des données collectées par les capteurs (*e.g.*, accéléromètre, robot de traite). En effet, la collecte automatisée des données entraîne encore actuellement un certain nombre de défis notamment dans le nettoyage des données, comme la gestion des données manquantes suite à une panne de l'automate ou à un défaut de connexion des capteurs, par exemple. A notre connaissance, aucune règles précises n'est défini quant à la correction des données manquantes ou aberrantes (méthodes à utiliser, nombre de données manquantes successives autorisées pour interpoler, seuil définissant une données aberrantes selon la variable étudiée). La question des données aberrantes est d'autant plus importante dans le contexte de la détection de perturbation. En effet, un point particulièrement haut ou bas peut être dû à une erreur de mesure par les technologies comme à une perturbation subit par l'animal. Les méthodes utilisées dans cette étude étant basées sur des lissages de données, elles permettent la prise en charge de petites interruptions dans la continuité des relevés. Cependant, un manque trop important risque d'entraîner des lissages abusifs et la perte des informations liés aux événements qui se sont déroulés pendant ce laps de temps.

La performance des capteurs eux-mêmes, comme les accéléromètres, est déterminante pour l'interprétation des résultats. En effet, une précédente publication a remis en question la fiabilité de la détection de certaines variables de l'accéléromètre utilisé dans cette étude, en particulier la rumination, en montrant une faible corrélation de celle-ci avec des observations effectuées par un observateur expérimenté (Bizeul *et al.*, 2022). Le temps passé debout semble bien mesuré par différents types d'accéléromètres (Brochers *et al.*, 2016), mais à notre connaissance, la fiabilité de la détection de ce comportement par le capteur Medria n'a pas été précisément étudiée.

Il semble donc pertinent de mener des études supplémentaires afin de valider la précision de la détection des comportements mesurés par ces technologies, mais aussi définir des règles permettant de prendre en compte les éventuels problèmes techniques (*e.g.*, données aberrantes ou manquantes) pouvant survenir avant d'utiliser les méthodes de détection en vue de leur application en élevage commercial.

4. Constitution du gold standard

La qualité du gold standard contenant les événements observés par les animaliers est également un élément crucial pour la fiabilité et l'interprétation des calculs de sensibilités et de spécificités. Certains facteurs pouvant influencer les variables étudiées n'ont pas été pris en compte dans le gold standard utilisé dans cette étude. Par exemple, la température qui peut entraîner un stress thermique et augmenter le temps que la vache passe debout (Mattachini *et al.*, 2017). D'autres facteurs peuvent avoir un impact sur ces variables, comme le statut hiérarchique de la vache qui peut influencer son budget-temps et la répartition de ses activités tout au long de la journée (Ungerfeld *et al.*, 2013) ainsi que son rendement laitier : les dominantes ont une production de lait plus importante que les dominées qui serait lié à un meilleur accès aux ressources alimentaires (Sołtysiak & Nogalski, 2010).

Il ne faut pas non plus exclure l'absence de certains événements liée à une erreur humaine, notamment concernant les événements sanitaires qui ne faisaient pas l'objet d'une observation autre que le suivi effectué habituellement. Ainsi, certains problèmes de santé ont pu ne pas être identifiés par les animaliers, et donc ne pas figurer dans le gold standard, conduisant à l'identification de faux négatifs si une perturbation est détectée par les méthodes.

Conclusion

L'objectif de cette étude était de voir si l'utilisation de technologies de l'élevage de précision combinées à des méthodes de lissage de données permettait d'identifier des perturbations dans la trajectoire de variables zootechniques et comportementales récoltées automatiquement par des capteurs et des automates utilisés dans les élevages laitiers. Les méthodes utilisées permettent de détecter des perturbations dans la trajectoire de ces variables. Cependant, d'autres études sont encore nécessaires afin d'améliorer leur performance et leur capacité à détecter des perturbations liées à des événements pouvant impacter le bien-être, la santé ou la production des animaux d'élevage, notamment en précisant les paramètres d'utilisation des méthodes (choix du degré de lissage, détermination du seuil de caractérisation d'une perturbation, choix des variables) pour les rendre plus pertinentes et fiables en vue d'une utilisation par les éleveurs.

Bibliographie

- Anses. (2018). *Avis de l'Agence nationale de sécurité sanitaire de l'alimentation, de l'environnement et du travail. Saisine n° « 2016-SA-0288 ». Bien-être animal : contexte, définition et évaluation*. 34. <https://www.anses.fr/fr/system/files/SABA2016SA0288.pdf>
- Barker, Z. E., Vázquez Diosdado, J. A., Codling, E. A., Bell, N. J., Hodges, H. R., Croft, D. P., & Amory, J. R. (2018). Use of novel sensors combining local positioning and acceleration to measure feeding behavior differences associated with lameness in dairy cattle. *Journal of Dairy Science*, *101*(7), 6310–6321. <https://doi.org/10.3168/jds.2016-12172>
- Bates, D., Maechler, M., & Bolker, B. (2024). *Linear Mixed-Effects Models using “Eigen” and S4* (1.1-35.3). <https://cran.r-project.org/web/packages/lme4/lme4.pdf>
- Ben Abdelkrim, A., Tribout, T., Martin, O., Boichard, D., Ducrocq, V., & Friggens, N. C. (2021). Exploring simultaneous perturbation profiles in milk yield and body weight reveals a diversity of animal responses and new opportunities to identify resilience proxies. *Journal of Dairy Science*, *104*(1), 459–470. <https://doi.org/10.3168/jds.2020-18537>
- Bizeul, N., Dufrasne, I., & Lessire, F. (2022, June 26). *Test of a sensor to estimate grazing and rumination time in dairy cow behaviour at pasture*. EGF. <https://orbi.uliege.be/handle/2268/306143>
- Blokhuis, H. J., Veissier, I., Miele, M., & Jones, B. (2010). The Welfare Quality® project and beyond: Safeguarding farm animal well-being. *Acta Agriculturae Scandinavica, Section A - Animal Science*, *60*(3), 129–140. <https://doi.org/10.1080/09064702.2010.523480>
- Borchers, M. R., Chang, Y. M., Tsai, I. C., Wadsworth, B. A., & Bewley, J. M. (2016). A validation of technologies monitoring dairy cow feeding, ruminating, and lying behaviors. *Journal of Dairy Science*, *99*(9), 7458–7466. <https://doi.org/10.3168/jds.2015-10843>
- Chu, K. (1999). An introduction to sensitivity, specificity, predictive values and likelihood ratios. *Emergency Medicine*, *11*(3), 175–181. <https://doi.org/10.1046/j.1442-2026.1999.00041.x>
- Cleveland, W. S., Grosse, E., & Shyu, W. M. (1992). Local Regression Models. In *Statistical Models in S* (1st ed., p. 68). Routledge.
- Cockburn, M. (2020). Review: Application and Prospective Discussion of Machine Learning for the Management of Dairy Farms. *Animals*, *10*(9), 1690. <https://doi.org/10.3390/ani10091690>
- Codrea, M. C., Højsgaard, S., & Friggens, N. C. (2011). Differential smoothing of time-series measurements to identify disturbances in performance and quantify animal response characteristics: An example using milk yield profiles in dairy cows1. *Journal of Animal Science*, *89*(10), 3089–3098. <https://doi.org/10.2527/jas.2010-3753>
- Corazzin, M., Romanzin, A., Foletto, V., Fabro, C., Da Borso, F., Baldini, M., Bovolenta, S., & Piasentier, E. (2021). Heat stress and feeding behaviour of dairy cows in late lactation. *Italian Journal of Animal Science*, *20*(1), 600–610. <https://doi.org/10.1080/1828051X.2021.1903818>
- David, I., Le, V., & Rohmer, T. (2023). *UpDown -An R Package to Identify and Characterize Disturbances from Longitudinal Observations*. <https://hal.inrae.fr/hal-04246616>

Durand, M., Largouët, C., De Beaufort, L. B., Dourmad, J.-Y., & Gaillard, C. (2023). Estimation of gestating sows' welfare status based on machine learning methods and behavioral data. *Scientific Reports*, *13*(1), 21042. <https://doi.org/10.1038/s41598-023-46925-z>

European Commission. (2023, October 10). *Eurobarometer shows how important Animal Welfare is for Europeans* [Text]. <https://ec.europa.eu/commission/presscorner/home/en>

Farm Animal Welfare Council. (2009). FAWC Report on Farm Animal Welfare in Great Britain: Past, Present and Future. *GOV.UK*. <https://www.gov.uk/government/publications/fawc-report-on-farm-animal-welfare-in-great-britain-past-present-and-future>

Gomez, A., & Cook, N. B. (2010). Time budgets of lactating dairy cattle in commercial freestall herds. *Journal of Dairy Science*, *93*(12), 5772–5781. <https://doi.org/10.3168/jds.2010-3436>

González, L. A., Tolkamp, B. J., Coffey, M. P., Ferret, A., & Kyriazakis, I. (2008). Changes in Feeding Behavior as Possible Indicators for the Automatic Monitoring of Health Disorders in Dairy Cows. *Journal of Dairy Science*, *91*(3), 1017–1028. <https://doi.org/10.3168/jds.2007-0530>

Green, L. E., Hedges, V. J., Schukken, Y. H., Blowey, R. W., & Packington, A. J. (2002). The Impact of Clinical Lameness on the Milk Yield of Dairy Cows. *Journal of Dairy Science*, *85*(9), 2250–2256. [https://doi.org/10.3168/jds.S0022-0302\(02\)74304-X](https://doi.org/10.3168/jds.S0022-0302(02)74304-X)

Le, V. (2022). *Nouvelle mesure de la robustesse des animaux d'élevage par utilisation des données de phénotypage haut-débit* [Phdthesis, INPT Toulouse]. <https://hal.inrae.fr/tel-03967884>

Le, V., Rohmer, T., & David, I. (2024). Identification and characterization of unknown disturbances in a structured population using high-throughput phenotyping data and measurement of robustness: application to growing pigs. *Journal of Animal Science*. <https://doi.org/10.1093/jas/skae059>

Levitin, D. J., Nuzzo, R. L., Vines, B. W., & Ramsay, J. O. (2007). Introduction to functional data analysis. *Canadian Psychology / Psychologie Canadienne*, *48*(3), 135–155. <https://doi.org/10.1037/cp2007014>

Mattachini, G., Antler, A., Riva, E., Arbel, A., & Provolo, G. (2013). Automated measurement of lying behavior for monitoring the comfort and welfare of lactating dairy cows. *Livestock Science*, *158*(1), 145–150. <https://doi.org/10.1016/j.livsci.2013.10.014>

Mattachini, G., Bava, L., Sandrucci, A., Tamburini, A., Riva, E., & Provolo, G. (2017). Effects of feed delivery frequency in different environmental conditions on time budget of lactating dairy cows. *Journal of Dairy Research*, *84*(3), 272–279. <https://doi.org/10.1017/S0022029917000310>

Mellor, D. J. (2017). Operational Details of the Five Domains Model and Its Key Applications to the Assessment and Management of Animal Welfare. *Animals: An Open Access Journal from MDPI*, *7*(8), 60. <https://doi.org/10.3390/ani7080060>

Nguyen-Ba, H., van Milgen, J., & Taghipoor, M. (2020). A procedure to quantify the feed intake response of growing pigs to perturbations. *Animal: An International Journal of Animal Bioscience*, *14*(2), 253–260. <https://doi.org/10.1017/S1751731119001976>

R core Team. (2023). *R: A Language and Environment for Statistical Computing* (4.3.2). R Foundation for Statistical Computing. <https://www.r-project.org/>

- Ramsay, J., Hooker, G., & Graves, S. (2024). *fda: Functional Data Analysis* (6.1.8). <https://cran.r-project.org/web/packages/fda/index.html>
- Reimert, I., Webb, L. E., van Marwijk, M. A., & Bolhuis, J. E. (2023). Review: Towards an integrated concept of animal welfare. *Animal*, *17*, 100838. <https://doi.org/10.1016/j.animal.2023.100838>
- Reith, S., Pries, M., Verhulsdonk, C., Brandt, H., & Hoy, S. (2014). Influence of estrus on dry matter intake, water intake and BW of dairy cows. *Animal*, *8*(5), 748–753. <https://doi.org/10.1017/S1751731114000494>
- Rohmer, T., Le, V., & David, I. (2023). *UpDown: Detecting Group Disturbances from Longitudinal Observations* (1.2.1). <https://cran.r-project.org/web/packages/UpDown/index.html>
- Sauvant, D., & Martin, O. (2010). Robustesse, rusticité, flexibilité, plasticité... les nouveaux critères de qualité des animaux et des systèmes d'élevage : définitions systémique et biologique des différents concepts. *INRAE Productions Animales*, *23*(1), 5–10. <https://doi.org/10.20870/productions-animales.2010.23.1.3280>
- Sinclair, M., Lee, N. Y. P., Hötzel, M. J., de Luna, M. C. T., Sharma, A., Idris, M., Derkley, T., Li, C., Islam, M. A., Iyasere, O. S., Navarro, G., Ahmed, A. A., Khruapradab, C., Curry, M., Burns, G. L., & Marchant, J. N. (2022). International perceptions of animals and the importance of their welfare. *Frontiers in Animal Science*, *3*. <https://doi.org/10.3389/fanim.2022.960379>
- Sitkowska, B., Kolenda, M., & Piwczyński, D. (2020). Comparison of the fit of automatic milking system and test-day records with the use of lactation curves. *Asian-Australasian Journal of Animal Sciences*, *33*(3), 408–415. <https://doi.org/10.5713/ajas.19.0190>
- Sołtysiak, T., & Nogalski, Z. (2010). The Effects of Social Hierarchy in a Dairy Cattle Herd on Milk Yield. *Polish Journal of Natural Science*, *25*(1), 22–30. <https://doi.org/10.2478/v10020-010-0002-1>
- Stygar, A. H., Frondelius, L., Berteselli, G. V., Gómez, Y., Canali, E., Niemi, J. K., Llonch, P., & Pastell, M. (2023). Measuring dairy cow welfare with real-time sensor-based data and farm records: a concept study. *Animal*, *17*(12), 101023. <https://doi.org/10.1016/j.animal.2023.101023>
- Taghipoor, M., Pastell, M., Martin, O., Nguyen Ba, H., van Milgen, J., Doeschl-Wilson, A., Loncke, C., Friggens, N. C., Puillet, L., & Muñoz-Tamayo, R. (2023). Animal board invited review: Quantification of resilience in farm animals. *Animal*, *17*(9), 100925. <https://doi.org/10.1016/j.animal.2023.100925>
- Ungerfeld, R., Cajarville, C., Rosas, M., & Repetto, J. (2014). Time budget differences of high- and low-social rank grazing dairy cows. *New Zealand Journal of Agricultural Research*, *57*(2), 122–127. <https://doi.org/10.1080/00288233.2014.893892>
- Veissier, I., Kling-Eveillard, F., Mialon, M.-M., Silberberg, M., de Boyer Des Roches, A., Terlouw, C., Ledoux, D., Meunier, B., & Hostiou, N. (2019). Élevage de précision et bien-être en élevage : la révolution numérique de l'agriculture permettra-t-elle de prendre en compte les besoins des animaux et des éleveurs? *INRA Productions Animales*, *32*(2), 281–290. <https://doi.org/10.20870/productions-animales.2019.32.2.2478>
- Wickham, H., François, R., Henry, L., Müller, K., Vaughan, D., Software, P., & PBC. (2023). *dplyr: A Grammar of Data Manipulation* (1.1.4). <https://cran.r-project.org/web/packages/dplyr/index.html>

Annexe

Caractérisation des perturbations selon le type d'événement

[résultats complémentaires ajoutés après soumission du rapport au jury de Master 2]

L'objectif de ces analyses complémentaires était de déterminer si les perturbations présentes dans la trajectoire des variables zootechniques et comportementales étudiées dans ce rapport possèdent des caractéristiques associées à un type d'événement particulier (*i.e.*, alimentaire, reproduction, sanitaire).

Pour répondre à cette question, il a été choisi d'utiliser les perturbations détectées par la méthode des B-splines (FDA) car elle permet d'identifier tout changement dans la trajectoire des données contrairement à UpDown qui ne fournit que les informations de certaines variations que la méthode identifie comme une perturbation. De même, aucun seuil de sélection basé sur l'importance des perturbations n'a été appliqué sur FDA afin de ne pas introduire de biais en éliminant les plus petits éléments par exemple. Ainsi, les analyses ont été faites sur l'ensemble des variations identifiées sur chaque vache pour les trois variables (production laitière, ingestion et temps passé debout). L'analyse a porté sur 4 caractéristiques dont les informations sont recueillies après l'application des lissages ($\lambda = 2,5$ et $\lambda = 10^4$) : la durée de la perturbation, le maximum et le minimum de différence entre les deux courbes de lissage et l'intensité de la perturbation (*i.e.*, différence entre le maximum et le minimum de différence entre les courbes). Des modèles linéaires mixtes ont été appliqués sur ces données après vérification des conditions d'application. Les données ont été transformées (sqrt pour la durée et log pour les autres). Les individus ont été définis comme facteurs aléatoires et le type d'événement (sanitaire, reproduction, alimentaire) ainsi que la variable sont définis comme facteurs fixes.

Les analyses ne montraient pas d'effet du type d'événement sur la durée des perturbations détectées sur les 3 variables. En revanche, une interaction significative a été observée entre le type d'événement et la variable pour le maximum de différence entre les courbes de lissage (LMM type 3, $p = 0,001$, Tableau A). En effet, cette caractéristique était plus importante pour les perturbations liées à l'alimentation que celles liées à la santé pour la variable comportementale (Test de Tukey, $p = 0,01$). De même, le maximum de différence entre les courbes observé sur la MSI est plus important pour les perturbations liées à un événement d'alimentation que de reproduction (Test de Tukey, $p = 0,02$). La durée moyenne d'une perturbation est plus longue sur la production laitière ou l'ingestion que sur le temps passé debout quel que soit le type d'événement (PL : LMM type 3, $p < 0,01$; MSI : LMM type 3, $p < 0,01$).

Inversement, l'intensité de la perturbation est plus forte pour la variable comportementale que pour la production laitière ou l'ingestion quel que soit le type d'événement (LMM type 3, $p < 0,01$). Aucun effet significatif de la variable ni du type d'événement n'a été relevé pour le minimum de différence entre les courbes de lissage, quelque soit la variable étudiée.

Bien que ces résultats ne montrent pas d'effet marqué du type d'événement sur les caractéristiques des perturbations des variables zootechniques, le maximum de différence entre les courbes semble présenter un intérêt pour la caractérisation des événements alimentaires, sanitaires et liés à la reproduction sur la variable comportementale. D'autre part, il semble que les perturbations identifiées sur le temps passé debout sont plutôt courtes et intenses quand celles identifiées sur la production laitière ou l'ingestion sont plus longues et d'intensité plus faible.

Il est toutefois nécessaire de compléter et confirmer ces résultats. Il serait pertinent notamment d'explorer davantage d'indicateurs pouvant caractériser une perturbation (*e.g.*, le délai entre l'événement et le début de la perturbation) et tester l'influence du type d'événement sur celui-ci. Il serait également pertinent d'étudier plus précisément les caractéristiques des jours précédents les perturbations détectées et liées à un événement observé par les animaliers afin de déterminer s'il est possible d'effectuer une détection a priori des perturbations. Prédire une perturbation, voire en identifier la cause, pourrait faciliter et optimiser les interventions sur les animaux en élevage, notamment les interventions vétérinaires.

Tableau A. Résultats des modèles linéaires mixtes (LMM) appliqués aux différentes variables (production laitière, la quantité de matière sèche ingérée et le temps passé debout) pour tester l'effet du type d'événement (alimentaire, sanitaire, reproduction) et son interaction avec les variables sur les différentes caractéristiques des perturbations.

| Caractéristiques | Type d'événement | Variables | Interaction |
|---|-------------------------|------------------|--------------------|
| Durée de la perturbation | p = 0,85 | p < 0,001 | p = 0,40 |
| Maximum de différence entre les courbes de lissage | p = 0,01 | p < 0,001 | p = 0,001 |
| Minimum de différence entre les courbes de lissage | p = 0,37 | p = 0,27 | p = 0,53 |
| Intensité de la perturbation | p = 0,01 | p < 0,001 | p < 0,01 |

Détection de ruptures dans la trajectoire des données de production et de comportement chez la vache laitière - Mise en relation avec les événements sanitaires, de reproduction et d'alimentation du troupeau

Résumé

Le bien-être animal est un enjeu d'actualité sur le plan scientifique, éthique, sociétal et commercial. En élevage bovin laitier, évaluer le bien-être des animaux reste cependant difficile en routine dans les fermes commerciales malgré l'environnement challengeant qu'elles constituent et des potentielles répercussions sur le bien-être et les performances de production des animaux. Dans cette étude, nous comparons deux méthodes de lissages (les B-splines et un lissage à noyau) appliquées à des données de production laitière, de quantité de matière sèche ingérée et au temps passé debout collectés par des capteurs et des automates afin de déterminer si elles permettent de détecter des perturbations liées à des événements préalablement référencés par les animaliers (*i.e.*, événements sanitaires, alimentaires et de reproduction). Les analyses montrent que le lissage à noyaux est très spécifique et que les B-splines sont plus sensibles que le lissage à noyaux. Il est toutefois nécessaire d'affiner la détermination des paramètres de détection (seuil de sélection des perturbations, degré de lissage), ainsi que le choix des variables sur lesquelles sont appliquées ces méthodes.

Mots clés : Automates, Bien-être animal, Elevage de précision, Lissage de données

Detection of disturbances in the trajectory of production and behavioral data for dairy cows - Link to herd health, reproduction and feeding events

Abstract

Animal welfare is a scientific, ethical, societal, and commercial issue. In dairy breeding, assessing animal welfare remains difficult on a routine basis on commercial farms, despite the challenging environment they represent and the potential repercussions on animal welfare and production performance. In this study, we compare two methods of smoothing data (B-splines and kernel smoothing) applied to milk yield, dry matter intake, and standing time collected by sensors and automatons to determine if it enables to detect disturbances linked to events previously referenced by animal handlers (*i.e.*, sanitary, feeding and reproductive event). Analysis shows that kernel smoothing is highly specific and that B-splines are more sensitive than kernel smoothing. However, further refinement is needed in determining the detection parameters (threshold for selecting potential disturbances, degree of smoothing), as well as the choice of variables on which these methods are applied.

Keys words: Animal welfare, Automatons, Data smoothing, Precision breeding