

L'Institut Agro Rennes-Angers
 Site d'Angers Site de Rennes

Année universitaire : 2021 - 2022 Spécialisation : Science des données	Mémoire de fin d'études <input type="checkbox"/> d'ingénieur de l'Institut Agro Rennes-Angers (Institut national d'enseignement supérieur pour l'agriculture, l'alimentation et l'environnement) <input checked="" type="checkbox"/> de master de l'Institut Agro Rennes-Angers (Institut national d'enseignement supérieur pour l'agriculture, l'alimentation et l'environnement) <input type="checkbox"/> de l'Institut Agro Montpellier (étudiant arrivé en M2) <input type="checkbox"/> d'un autre établissement (étudiant arrivé en M2)
---	---

Amélioration d'un outil de détection d'évènements (techniques ou sanitaires) à l'échelle de plusieurs bandes de truies

Par : Gaëlle HAMELIN



Soutenu à Rennes le 30 Août 2022

Devant le jury composé de :

Président : David Causeur PR Institut Agro

Autre membre du jury : Mathieu Emily PR Institut Agro

Maîtres de stage : Charlotte Gaillard CR-HDR INRAE

Tom Rohmer CR INRAE

Organisme d'accueil : INRAE

Enseignant référent : David Causeur

Les analyses et les conclusions de ce travail d'étudiant n'engagent que la responsabilité de son auteur et non celle de l'Institut Agro Rennes-Angers

Confidentialité

Non Oui si oui : 1 an 5 ans 10 ans

Pendant toute la durée de confidentialité, aucune diffusion du mémoire n'est possible ⁽¹⁾.

Date et signature du maître de stage ⁽²⁾ : 27/07/2022



A la fin de la période de confidentialité, sa diffusion est soumise aux règles ci-dessous (droits d'auteur et autorisation de diffusion par l'enseignant à renseigner).

Droits d'auteur

L'auteur⁽³⁾ **Hamelin Gaëlle**

autorise la diffusion de son travail (immédiatement ou à la fin de la période de confidentialité)

Oui Non

Si oui, il autorise

la diffusion papier du mémoire uniquement⁽⁴⁾

la diffusion papier du mémoire et la diffusion électronique du résumé

la diffusion papier et électronique du mémoire (joindre dans ce cas la fiche de conformité du mémoire numérique et le contrat de diffusion)

(Facultatif) accepte de placer son mémoire sous licence Creative commons CC-By-Nc-Nd (voir Guide du mémoire Chap 1.4 page 6)

Date et signature de l'auteur : 09/08/22



Autorisation de diffusion par le responsable de spécialisation ou son représentant

L'enseignant juge le mémoire de qualité suffisante pour être diffusé (immédiatement ou à la fin de la période de confidentialité)

Oui Non

Si non, seul le titre du mémoire apparaîtra dans les bases de données.

Si oui, il autorise

la diffusion papier du mémoire uniquement⁽⁴⁾

la diffusion papier du mémoire et la diffusion électronique du résumé

la diffusion papier et électronique du mémoire

Date et signature de l'enseignant :

Remerciements

Je tiens tout d'abord à remercier l'Institut Agro Rennes-Angers de m'avoir soutenue et accompagnée dans la réalisation de ma formation diplômante de Master 2 en Sciences des données.

Je voudrais ensuite remercier les membres de l'équipe enseignante de l'Unité Pédagogique de mathématiques appliquées de m'avoir fait confiance et de m'avoir transmis autant de nouvelles connaissances qui me seront très utiles dans la suite de mon parcours professionnel.

Je remercie en particulier David Causeur, mon professeur référent, pour son suivi, sa disponibilité et son accompagnement dans mes projets.

Je remercie également les membres de l'Unité Pédagogique des Sciences du sol de m'avoir soutenue dans ce beau projet d'évolution professionnelle.

Surtout, j'adresse mes sincères remerciements à mes encadrants de stage, Charlotte Gaillard et Tom Rohmer, tous deux chercheurs à l'Inrae qui m'ont accompagnée, orientée et conseillée durant toute la durée de mon stage. J'ai beaucoup appris auprès d'eux et j'en sors grandie et enrichie !

Je remercie aussi les personnes de l'UMR Pégase à Saint-Gilles avec qui j'ai pu échanger et passer de bons moments (Anne, Maeva, Rémy, Ellynn, Myriam, Chantal, Laetitia, Carole, Patrick, Marie, Malinka, Clémentine, Chloé, Emilie...)

Enfin, je tiens tout particulièrement à remercier mon conjoint pour son soutien inconditionnel et ses encouragements qui m'ont portée pendant un an.

Table des matières

1. Introduction	1
1.1. Contexte.....	1
1.2. Objectifs et enjeux	2
1.3. Hypothèses	2
2. Matériel et méthode	3
2.1. Conditions expérimentales	3
2.2. Analyse des données	4
2.2.1. Méthode de lissage différentiel	5
2.2.2. Tests non paramétriques de détection de ruptures.....	6
2.2.3. Random Forest.....	7
2.2.4. Evaluation de la performance des différentes méthodes	8
3. Résultats et discussion	8
3.1. Analyse descriptive des données	8
3.2. Evènements perturbateurs	11
3.2.1. Evènement thermique Chaud	12
3.2.2. Evènement thermique Froid.....	12
3.2.3. Evènement Alimentaire	12
3.2.4. Evènement Malade	13
3.3. Méthode de lissage différentiel	14
3.4. Tests non paramétriques de détection de ruptures de type <i>CUSUM</i>	17
3.5. Random Forest	18
4. Conclusion et perspectives	19

Table des figures

Figure 1 : Design expérimental sur les bandes de truies gestantes selon les jours de gestation.....	3
Figure 2 : Schéma représentatif de l'aire, de la distance maximum et de la distance moyenne entre les 2 courbes de lissage différent	5
Figure 3 : Moyenne par truie et par jour de gestation en situation basale pour les (3.a) Nombre de buvées, (3.b) Quantité d'eau bue (L), (3.c) Temps passé allongé (min) (3.d) Durée d'abreuvement (min)	10
Figure 4 : Variations intra-individuelles de la durée d'abreuvement journalier pour 2 truies de la bande 17.....	11
Figure 5 : Boxplots de la quantité d'eau bue et du temps passé allongé selon l'Evènement...11	
Figure 6 : Fréquence d'abreuvement d'une truie de la bande 6 malade pendant la gestation.13	
Figure 7 : (7.a) Temps passé debout pour une truie de la bande 16 pendant la gestation (7.b) Quantité d'eau bue pour une truie de la bande 5 pendant la gestation	14

Table des Tableaux

Tableau 1 : Choix des hyperparamètres.....	7
Tableau 2 : Matrice de confusion.....	8
Tableau 3 : Profils d'évolution des variables par rapport à la situation basale selon l'évènement.....	14
Tableau 4 : Mesure de la performance des tests par variable tous Evènements confondus pour le quantile 0.95 (a) et le quantile 0.85 (b) (méthode du lissage différentiel).....	15
Tableau 5 : Mesure de la performance des tests par Bande selon l'Evènement pour le quantile 0.95(a) et le quantile 0.85 (b) (méthode du lissage différentiel), toutes bandes confondues.....	16
Tableau 6 : Mesure de la performance des tests par Bande selon l'Evènement (méthode de détection de ruptures).....	18
Tableau 7 : Mesure de la performance des tests par Bande selon l'Evènement (Random Forest).....	19

Liste des abréviations

UMR PEGASE : Unité Mixte de Recherche Physiologie, Environnement et Génétique pour l'Animal et les Systèmes d'Élevage

UE3P : Unité Expérimentale Physiologie et Phénotypage des Porcs

DAC : Distributeur Automatique de Concentrés alimentaires

CIE : Compteur Individuel d'Eau

AC : Abreuvoirs Connectés

OAD : Outil d'Aide à la Décision

RFID : Radio Frequency Identification

1. Introduction

1.1. Contexte

Il existe actuellement en France plus de 14 000 élevages de porcs. Le développement du numérique, dans un contexte d'élevage de précision, permet de proposer aux éleveurs des outils pour optimiser le suivi individualisé des animaux afin de favoriser la détection des perturbations techniques et les troubles de santé.

En effet, les événements techniques sont davantage susceptibles de se produire dans les élevages porcins qui sont conduits en conditions très standardisées (contrôle minutieux du taux d'humidité, de la température, du taux d'ammoniac...). Un événement technique indésirable peut correspondre à la panne d'un appareil de régulation d'ambiance d'élevage ou à des dysfonctionnements des systèmes d'alimentation en eau ou en concentrés alimentaires des animaux. Ces événements ne sont pas toujours faciles à détecter ou du moins rapidement. Cependant, ils peuvent occasionner des pertes d'effectifs et donc de production importantes car ils sont susceptibles de concerner tout le groupe, ce qui peut avoir des répercussions économiques désastreuses. On estime que la perte d'une truie pleine en fin de gestation coûte près de 1000 euros à l'éleveur (*Ifip*).

Les problèmes de santé ou de bien-être sont en effet difficiles à détecter dans les grands groupes de truies en élevage intensif et encore plus à l'échelle individuelle. Des outils de surveillance individuelle se sont développés ces dernières années. Ils consistent à équiper chaque truie de capteurs auriculaires contenant une puce RFID permettant de tracer les comportements individuels d'abreuvement et d'alimentation. D'autre part, il est possible d'équiper les truies d'accéléromètres auriculaires afin d'apprécier l'activité physique de chaque animal (Cornou et Lundbye-Christensen, 2012). Cependant, ces capteurs sont plus fragiles et coûteux. Ils sont de plus un peu plus volumineux et donc moins bien tolérés par les truies qui ont tendance à les enlever. Par ailleurs, le traitement des informations issues des accéléromètres nécessite d'effectuer des pré-calculs et un nettoyage des données qui ne permettent pas actuellement de proposer aux éleveurs une solution complète.

Les abreuvoirs connectés permettent de mettre en évidence des variations anormales dans la dynamique d'abreuvement individuelle des truies pouvant traduire une modification de leur état physiologique ou de leur environnement. Il a été montré que les porcs peuvent être plus ou moins motivés à boire selon l'intensité et la nature des troubles sanitaires qu'ils présentent (Junge et al., 2012). Le comportement d'abreuvement est également étroitement lié à la température ambiante dans laquelle évoluent les truies (Bird et al., 2001; Seddon et al., 2011).

Chez les porcs, le volume d'eau consommé semble être un indicateur d'intérêt pour bien comprendre leur état de santé (Rousselière et al., 2017). La fièvre induit généralement une diminution ou une perte d'appétit qui va faire baisser en conséquence la motivation à boire (Martineau et Morvan, 2010). Par ailleurs, chez les truies gestantes, la consommation d'eau augmente à mesure que le moment de la mise bas approche et surtout dans les derniers jours de gestation (Mateu Chaler, 2020). Le niveau basal est fluctuant selon le stade de la gestation.

De la même façon, l'utilisation de capteurs d'activité permet d'alerter sur un comportement anormal, synonyme d'un problème latent. Les accéléromètres permettent la détection des modifications du rythme d'activité, de la locomotion ou du comportement habituel individuel. On sait que le stade de gestation influence également l'activité des truies. Elles sont moins actives à mesure que la mise bas approche et elles restent plus allongées.

La surveillance des truies durant la période péri-natale par des capteurs est efficace et bien moins chronophage que par vidéo (Cornou et al, 2014).

Des travaux antérieurs ont permis de montrer qu'il était possible de détecter des événements techniques ou sanitaires survenant dans un élevage bovin en analysant les données d'abreuvement par une méthode de lissage différentiel (Gaillard et al., 2015 ; Codrea et al., 2011). Chez les truies gestantes, en conditions expérimentales, cette méthode a également permis de détecter des événements techniques indésirables à partir de données d'abreuvement (Nizzi et al., stage M2 2021). Cependant la sensibilité de la méthode était inférieure à 50%. Afin d'améliorer les performances de la méthode de détection par lissage différentiel, nous avons décidé d'ajouter à nos méthodes les données d'activité fournies par les accéléromètres. D'autre part, nous avons choisi de tester une autre méthode de détection basée sur les ruptures de corrélations entre les divers couples de variables (Kojadinovic et al, 2016), en supposant qu'un événement perturbateur peut se traduire par un changement dans la relation entre les différentes variables. Enfin, nous avons testé la méthode du Random Forest car une étude a montré que les forêts aléatoires permettent de prédire les boiteries de truies gestantes à partir des données issues du comportement alimentaire, hydrique et d'activité, un jour à l'avance avec une sensibilité de 69% et une spécificité de 93% (Courboulay et al., 2017 ; Marcon et al., 2020).

1.2. Objectifs et enjeux

L'analyse des données présentée dans ce rapport a pour objectif de répondre à plusieurs questions :

- Peut-on détecter des événements induits dans des séries temporelles de suivi des différentes variables associées au comportement d'abreuvement et d'activité des truies gestantes ?
- Les modifications constatées dans les courbes temporelles pour les différentes variables sont-elles relatives au type d'évènement perturbateur subi par l'animal ?
- Quelle méthode donne les meilleurs résultats ?

1.3. Hypothèses

Dans nos conditions expérimentales, nous supposons qu'il est possible de détecter et de quantifier pour chaque variable et pour chaque truie, des irrégularités dans les dynamiques d'abreuvement et d'activité physique en lien avec les événements induits.

Nous pensons que l'ajout dans nos modèles des variables liées à l'activité des animaux, devrait permettre une meilleure détection des événements perturbateurs et améliorer la précision des méthodes. D'autre part, nous émettons l'hypothèse que les relations entre les différentes variables peuvent changer en périodes d'évènements.

2. Matériel et méthode

2.1. Conditions expérimentales

Les données ont été collectées à l'UE3P de Saint Gilles, à partir de 63 truies gestantes de type Large White croisées Landrace, réparties en 4 bandes, sur 2 périodes différentes allant d'Août à Novembre 2020 puis de Janvier à Avril 2021 (Figure 1). Les truies d'une même bande ont été inséminées en même temps (± 2 jours près) et sont donc toutes au même stade de gestation.

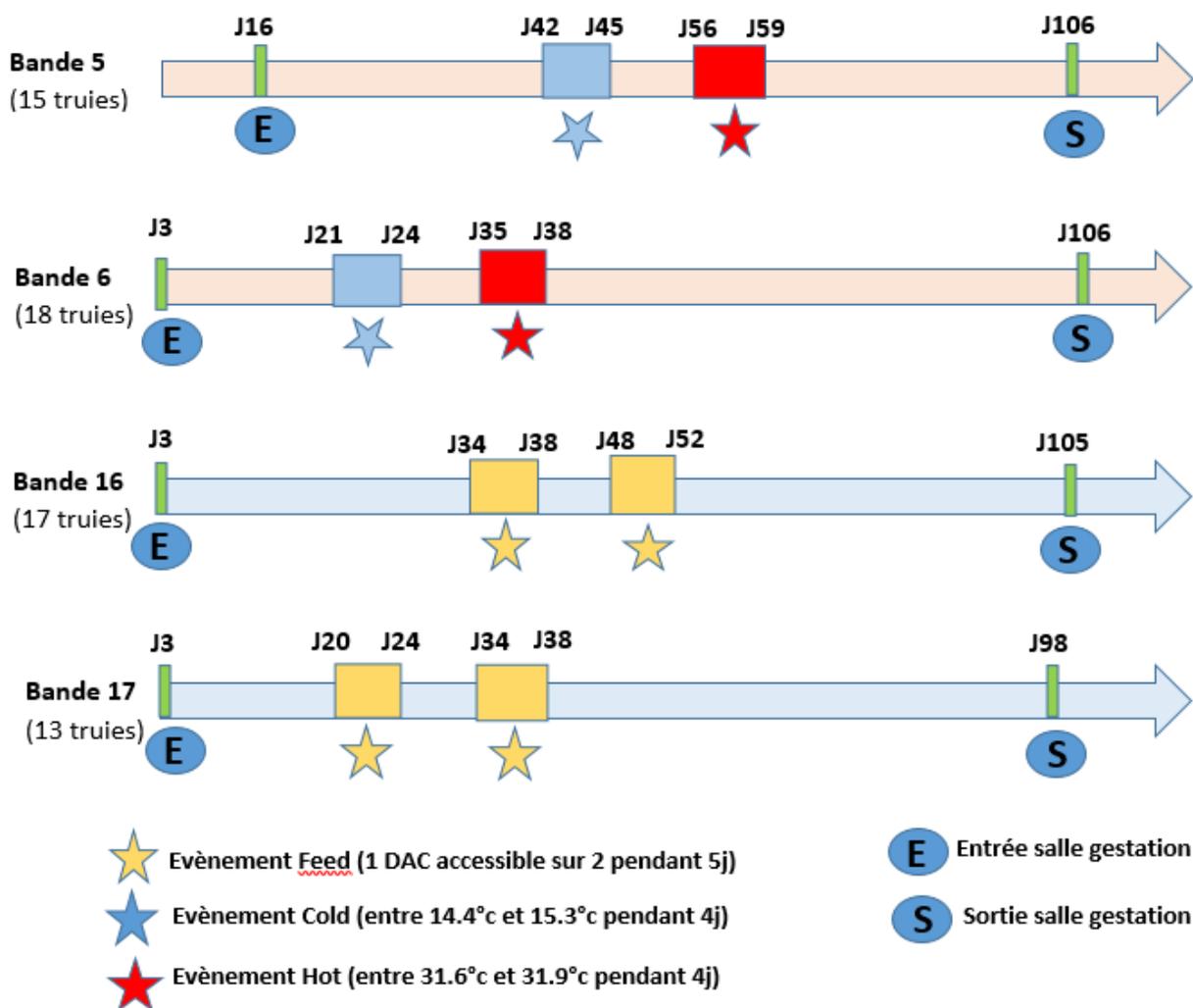


Figure 1 : Design expérimental sur les bandes de truies gestantes selon les jours de gestation

La durée moyenne de la gestation chez une truie est de 114 jours (3mois, 3 semaines et 3 jours). Les truies sont entrées en salle de gestation à J3 (3 jours après l'insémination) ou à J16 pour la bande 5. Elles ont quitté la salle de gestation à J98 ou J106, soit environ une semaine avant la mise-bas, pour aller ensuite en salle de maternité.

Chaque bande est logée dans une salle de gestation équipée de deux distributeurs automatiques de concentrés alimentaires (DAC) permettant la délivrance à chaque truie d'une ration individuelle de 2.9 kg/jour en moyenne (Gaillard et al, 2022) et de deux abreuvoirs connectés ASSERVA® (AC) encore appelés Compteurs Individuels d'Eau (CIE). Des antennes RFID permettent de détecter les passages individuels des truies aux abreuvoirs grâce à une lecture par radiofréquence des boucles auriculaires d'identification qu'elles portent. Ainsi, à chaque visite, le système collecte des données sur la quantité d'eau bue, la date et l'heure de la buvée ainsi que la durée d'abreuvement de chaque animal. Le calcul a permis de déterminer la fréquence journalière des buvées ainsi que la quantité d'eau ingérée et le temps passé à boire par truie chaque jour. Par ailleurs, chaque truie est également équipée d'un accéléromètre auriculaire (RFtrack, Rennes, France) qui enregistre une donnée par minute et transmet ces données qui sont agrégées chaque heure par un logiciel en données comportementales. Il y avait de nombreuses données manquantes concernant les relevés issus des accéléromètres. En effet, certaines truies ont perdu leur accéléromètre durant plusieurs heures ou il y a eu des problèmes techniques de dysfonctionnement. Nous disposons ainsi pour chaque truie, sur les heures observées, de la somme en min du temps passé debout, allongé ou en mouvement et du nombre de changements de position.

Les truies des bandes 16 et 17 ont été soumises à 2 événements alimentaires successifs de 5 jours chacun espacés de 10 jours et durant lesquels un seul distributeur d'aliments sur les deux était accessible. Les truies des bandes 5 et 6 ont subi un événement thermique froid puis chaud durant 4 jours consécutifs et espacés de 10 jours. La température habituelle moyenne dans les salles de gestation est de 20°C (± 2). L'événement thermique froid a consisté à arrêter le chauffage dans la salle de gestation et à ouvrir les fenêtres, entraînant une baisse de la température qui est descendue entre 14,4 et 15,3 °C. L'événement thermique chaud a été obtenu en augmentant la température dans la salle de gestation entre 31,6 et 31,9 °C et en arrêtant la ventilation. (voir Figure 1)

Ces événements ont été induits pour mimer les aléas techniques pouvant survenir dans un élevage. Par ailleurs, quelques événements sanitaires (maladies, boiteries...) ont eu lieu pendant les expérimentations et ont été consignés dans un registre mentionnant l'ensemble des problèmes particuliers, les interventions humaines réalisées, la date et les animaux concernés.

2.2. Analyse des données

Face au manque de données d'activité sur certaines périodes et pendant plusieurs heures consécutives, nous avons décidé de déterminer pour chaque jour et chaque truie, la moyenne horaire du temps passé debout, allongé et du nombre de changements de position.

Les données d'abreuvement qui sont collectées à chaque passage d'une truie à l'abreuvoir (durée de chaque abreuvement et quantité d'eau bue lors de cette buvée) ont été sommées par jour et par truie.

A l'issue du pré-traitement des données brutes, le jeu de données nettoyé comporte 6316 lignes et 12 variables qui sont soit qualitatives : bande (5,6,16,17), type d'évènement (froid, chaud, alimentaire vs. basal), jour de gestation, soit quantitatives : quantité d'eau bue (en litre), nombre de buvées, durée d'abreuvement (min), temps moyen horaire passé allongé (min) ou debout (min) et nombre moyen horaire de changements de position. Chaque ligne du tableau de

données représente une truie par jour de gestation avec les valeurs correspondant aux différentes variables.

L'ensemble des analyses a été réalisé avec le logiciel R (R Core Team, 2021).

Une application R Shiny (disponible à la demande) a été réalisée afin de présenter le jeu de données, les courbes temporelles par truie et par variable ainsi que les analyses bivariées pour chaque couple de variables, par truie.

Les coefficients de variations (CV) ont été calculés pour les variables d'abreuvement par truie, afin d'estimer la variance intra-individuelle. Puis, la moyenne des CV individuels a permis de déterminer le CV moyen indicateur de la variance inter-individuelle.

Une analyse de la variance par des modèles linéaires à effets mixtes a été réalisée pour chacune des 6 variables quantitatives à expliquer (temps passé debout ou allongé, nombre de buvées ou de changements de position, quantité et durée des buvées) en fonction de l'effet « Bande » (variable qualitative aléatoire), de l'effet « Evènement » (variable qualitative fixe) et de l'interaction « Evènement / bande » (aléatoire). En effet, dans le cadre des expérimentations, les truies ont été réparties aléatoirement dans les bandes et les évènements perturbateurs induits ont été choisis et fixés au départ.

Quand il y avait un effet Bande, une régression linéaire a été réalisée pour cette variable et par bande pour regarder s'il y avait un effet évènement (Froid / Chaud ou Alim) significatif sur cette bande.

Afin de détecter les évènements induits sur les 4 bandes de truies (évènements thermiques chaud, froid ou alimentaires), trois méthodes de détection différentes ont été testées et comparées.

2.2.1. Méthode de lissage différentiel

La méthode de lissage différentiel est utilisée pour mettre en évidence les potentielles variations dues aux évènements perturbateurs, à partir des courbes temporelles de chaque variable et pour chaque truie.

Il s'agit d'une méthode d'analyse des données fonctionnelles (FDA) basée sur des b-splines. Elle est mise en oeuvre à l'aide du package R *fda* (Ramsay et al., 2010). Les fonctions splines sont construites en joignant des polynômes de degré 3 bout à bout à des valeurs d'argument appelées nœuds (4 dans notre étude). Le principe de la méthode consiste à obtenir des courbes avec des degrés de lissage différents en appliquant des pénalités plus ou moins fortes sur les dérivés des fonctions. Ainsi, pour chaque truie, et pour chaque variable, 2 pénalités différentes (forte = 100 et faible = 10) ont permis d'établir 2 courbes plus ou moins lissées.

L'aire entre les 2 courbes, le maximum de la différence et la moyenne entre les 2 courbes ont ensuite été calculées (*Figure 2*).

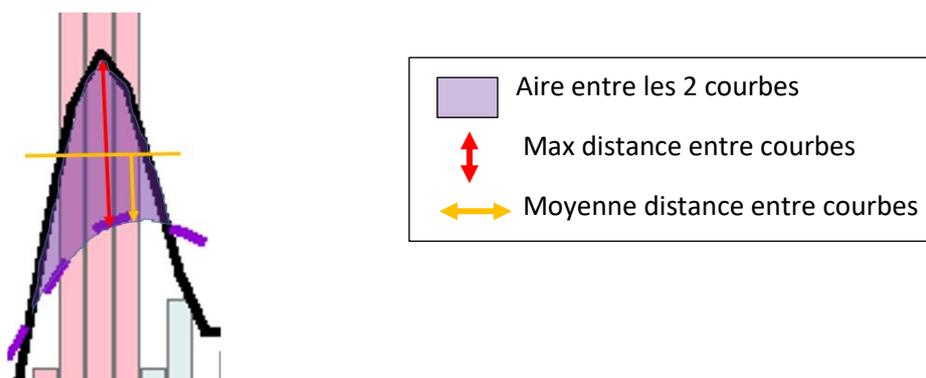


Figure 2 : Schéma représentatif de l'aire, de la distance maximum et de la distance moyenne entre les 2 courbes de lissage différent

Un seuil de positivité du test a été choisi à partir des 5% (quantile 0.95) ou des 15% (quantile 0.85) plus fortes valeurs (en valeur absolue) de chacune des 3 variables (aire entre les 2 courbes lissées, maximum de la différence, moyenne entre les 2). Nous avons ainsi attribué un 1 (predicted=positif) aux plus fortes valeurs absolues qui représentent les plus grandes différences en valeurs absolues entre les 2 courbes de lissage différent (augmentations ou diminutions). Un 0 a été attribué (predicted=négatif) pour les valeurs moins élevées des 3 variables.

En parallèle, nous avons affecté une valeur binaire de 1 (observed=positif), quand il s'agissait réellement d'un évènement induit. Nous avons affecté un 0 quand il n'y avait pas d'évènement (observed=négatif).

Des calculs de spécificité, sensibilité, précision, valeur prédictive positive, proportion d'évènements ont été réalisés pour évaluer la méthode de détection (*paragraphe 2.2.4*).

2.2.2. Tests non paramétriques de détection de ruptures

Des tests de détection de ruptures de corrélation entre 2 variables ont été utilisés afin d'estimer leur capacité à détecter les évènements thermiques ou alimentaires subis par chaque truie pendant la gestation. Nous faisons l'hypothèse que la façon dont interagissent les variables peut changer en période d'évènement perturbateur.

La méthode consiste à regarder si les corrélations de Spearman entre chaque couple de variables continues et pour chaque truie varient au cours d'un évènement.

Le coefficient de corrélation de Spearman oscille entre -1 et 1. Il peut s'écrire comme la corrélation de Pearson entre les rangs des observations et définit donc un indice de corrélation non-linéaire ; plus adéquat lorsque la relation entre les variables n'est pas nécessairement linéaire.

Nous avons réalisé des tests basés sur la théorie des tests « cumulative sum » par couple de variables et pour chaque truie, afin de repérer les éventuels points de changement au niveau des intervalles de temps considérés à l'aide du package R *npcp* (*Kojadinovic et al., 2020*).

Les intervalles de temps testés contenaient chacun un évènement :

- avant, pendant et après le 1^{er} évènement (Froid ou Alim1) c'est-à-dire du début des mesures (J3 ou J16) jusqu'à 7 jours après le 1^{er} évènement.
- un peu avant, pendant et après le 2^{ème} évènement, c'est-à-dire trois jours avant le 2^{ème} évènement (Chaud ou Alim2) et jusqu'à la fin des mesures.

La méthode est basée sur un ré-échantillonnage bootstrap (1000 répétitions réalisées) en considérant les observations non nécessairement indépendantes dans le temps.

Pour voir s'il y a un changement dans le coefficient de corrélation de Spearman sur les intervalles de temps, le test consiste à couper l'échantillon en 2 parties successives avec la taille des deux sous-échantillons qui varie puis à calculer les corrélations de Spearman sur chacun des deux sous-échantillons pour enfin regarder s'il y a une différence pondérée entre ces deux valeurs.

Sous l'hypothèse nulle (la corrélation de Spearman ne varie pas au cours du temps), la plus grande différence entre ces deux valeurs de Spearman va tendre à être faible, entraînant un non-rejet de l'hypothèse nulle.

Par contre, en présence d'un changement, la statistique de test va s'écarter de 0, entraînant un rejet de l'hypothèse nulle. L'estimation du point de changement (*ECP*) correspond à l'instant qui maximise la différence en valeur absolue entre les deux sous-échantillons.

Ainsi, lorsque le test rejette l'hypothèse nulle d'absence de changement dans la corrélation de Spearman (p -valeur < 0.05), nous avons estimé le point de changement correspondant à l'instant de rupture afin de vérifier s'il corrobore avec une période d'évènement.

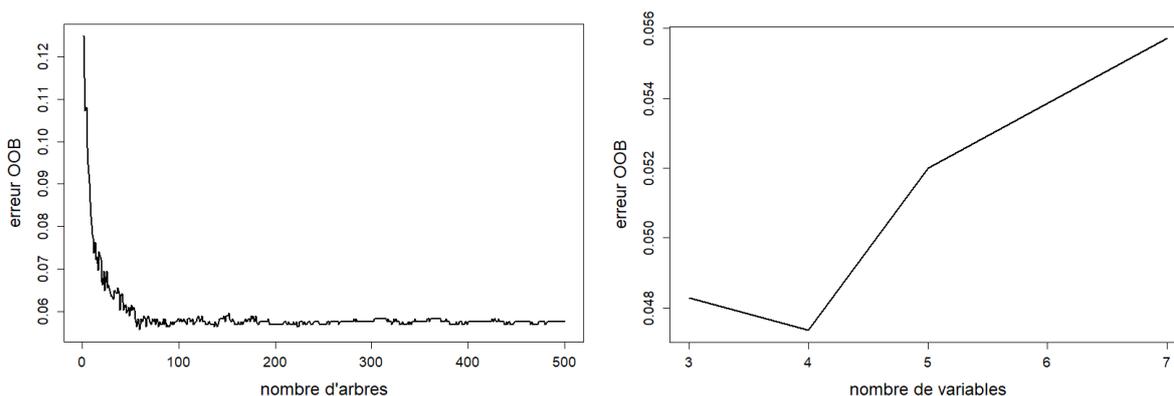
Dans un premier temps, pour chaque truie et chacun des couples de variables, nous avons attribué un 1 quand il y avait un ECP détecté par la méthode sur l'intervalle de temps analysé (predicted=positif) et un 0 en cas contraire (predicted=négatif). Puis, connaissant les jours réels des évènements perturbateurs, nous avons attribué un 1 (observed=positif) quand l'ECP détecté tombait juste pendant l'évènement et un 0 en cas contraire (observed=négatif). Dans un 2^{ème} temps, par truie et par intervalle de temps testé, les VP (vrais positifs) représentent le fait qu'au moins un des ECP détectés dans l'intervalle selon le couple de variables, corrobore avec l'évènement. Les FP (faux positifs) représentent le fait qu'au moins un des ECP détectés dans l'intervalle selon le couple de variables, ne corrobore pas avec l'évènement et les FN (faux négatifs) signifient qu'il n'y a pas eu d'ECP détecté dans l'intervalle quelque soit le couple de variables. Ainsi, par truie et par intervalle de temps testé, il est indiqué soit VP, soit FP soit FN. Il n'y avait pas de VN (vrais négatifs) car les intervalles de temps testés contenaient toujours un évènement.

Nous avons ensuite évalué la méthode de détection de ruptures de corrélations en calculant la sensibilité et la valeur prédictive positive (VPP).

2.2.3. Random Forest

L'analyse par forêts aléatoires est mise en œuvre à l'aide du package R randomForest (Andy Liaw et al 2002).

Etant donné la grande variabilité inter-individus, il aurait été préférable, pour une analyse plus fine, d'employer la méthode par truie mais nous ne disposons pas de suffisamment de données par truie (1 seule période de gestation avec évènements par truie). La méthode du Random Forest a donc été appliquée par bande de truies. La variable réponse à expliquer est la variable qualitative « Evènement » (classification). Les variables explicatives sont les variables quantitatives relatives au comportement d'abreuvement et à l'activité des truies. Pour chaque bande, nous avons défini les hyperparamètres « ntree » (nombre d'arbres dans le modèle avec tirage aléatoire des individus) et « mtry » (nombre de variables choisies à la construction de chaque nœud) permettant d'optimiser le taux d'erreur de mauvais classement (*Tableau 1*).



	Bande 5	Bande 6	Bande 16	Bande 17
ntree	300	200	200	300
mtry	4	4	4	4

Tableau 1 : Choix des hyperparamètres

Pour chaque analyse, 70% du jeu de données a été utilisé pour faire apprendre le modèle et 30% a été conservé pour tester le modèle. Différents indicateurs d'évaluation de la méthode ont été déterminés à partir des valeurs binaires (0 ou 1) de l'état prédit et celui observé.

2.2.4. Evaluation de la performance des différentes méthodes

Les évènements induits ont été associés (ou pas) aux perturbations détectées par les méthodes, ce qui a permis d'effectuer des calculs d'indicateurs visant à évaluer chaque méthode (sensibilité, spécificité, précision, valeur prédictive positive, taux d'évènements)

Pour les calculs, nous avons utilisé la classification binaire suivante (*Tableau 2*) :

		Predicted	
		0	1
Observed	0	VN	FP
	1	FN	VP

Tableau 2 : Matrice de confusion

VP = Vrai Positif (détecté positif par la méthode et réellement positif)

FN = Faux Négatif (détecté négatif par la méthode mais réellement positif)

VN = Vrai Négatif (détecté négatif par la méthode et réellement négatif)

FP = Faux Positif (détecté positif par la méthode mais réellement négatif)

Sensibilité = probabilité de détecter un Positif sachant qu'il est réellement Positif

$$\text{Sensibilité} = \text{VP} / (\text{VP} + \text{FN})$$

Spécificité = probabilité de détecter un Négatif sachant qu'il est réellement Négatif

$$\text{Spécificité} = \text{VN} / (\text{VN} + \text{FP})$$

$$\text{Précision} = (\text{VP} + \text{VN}) / (\text{VN} + \text{FP} + \text{FN} + \text{VN})$$

Taux de bonnes détections (Positives ou Négatives)

Valeur prédictive positive = probabilité d'être réellement Positif sachant qu'il est détecté Positif

$$\text{VPP} = \text{VP} / (\text{VP} + \text{FP})$$

Proportion d'évènements = nombre total d'évènements par rapport aux non évènements

3. Résultats et discussion

3.1. Analyse descriptive des données

Dans cette étude nous avons constaté que les truies en situation basale boivent en moyenne 8.3 litres/jour d'eau et qu'il existe une forte variation selon l'animal considéré (100 ml/jour à plus

de 60 litres/jour pour quelques individus atypiques). La variabilité inter-individuelle est importante. En effet, le coefficient de variation moyen toutes truies confondues pour la quantité d'eau bue par jour est de 38 %.

Effectivement, la consommation journalière en eau d'une truie gestante en bonne santé est très variable selon les individus. Ainsi, Klopfenstein *et al.* (1996) indiquent que la prise hydrique varie de 5 litres à plus de 30 litres/jour. Cerneau *et al.* (1997) évoquent une consommation journalière de 20 litres en moyenne. Une consommation quotidienne de plus de 20 litres devrait être mise en lien avec un comportement de potomanie dont la conséquence directe est l'augmentation importante des effluents de porcherie. Dans notre étude, trois truies ont présenté un comportement de potomanie pendant leur gestation (48 litres/jour).

On constate également que la quantité d'eau bue diminue au cours de la gestation, sauf à la fin où elle remonte (*Figure 3.b.*), de façon concomitante avec la durée d'abreuvement.

En moyenne sur toute la durée de la gestation, les truies se rendent 11 fois/jour à l'abreuvoir. La variabilité inter-individuelle est élevée car le CV moyen pour la fréquence journalière des buvées est de 41%.

La fréquence des buvées journalières diminue avec l'avancée de la gestation (*Figure 3.a*). On peut remarquer qu'au début de la gestation, les truies gravides se rendent en moyenne 19 fois/jour à l'abreuvoir contre 10 fois/ jour en fin de gestation. Ceci pourrait être mis en relation avec une limitation des déplacements pour éviter les pertes énergétiques et privilégier le repos.

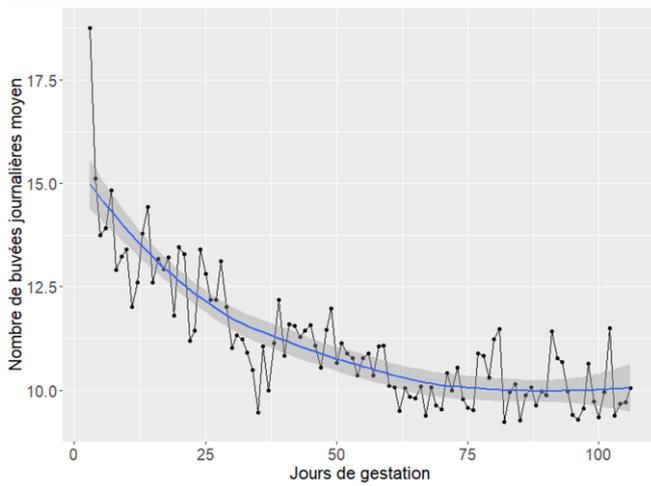
Ainsi, le temps passé allongé augmente avec la gestation (*Figure 3.c*). En début de gestation (3^{ème} Jour J3) une truie gravide passe en moyenne 72% de son temps couchée puis 80% en fin de gestation (J97). Comme mentionné par Cariolet et al (1985), l'activité des truies gestantes est moins importante au début et à la fin de la gestation. Elles passent en moyenne 46 min allongées par heure c'est à dire 77% du temps journalier et seulement 18% de leur temps debout et 5% en mouvement.

Pour compenser la diminution de leur mobilité, les truies boivent plus longtemps à chaque buvée quand la gestation avance (*Figure 3.d*), ce qui leur permet d'absorber suffisamment d'eau et de couvrir leurs besoins quotidiens. On constate une grande variabilité inter-individuelle puisque le CV moyen pour la durée des buvées journalière est de 40 %.

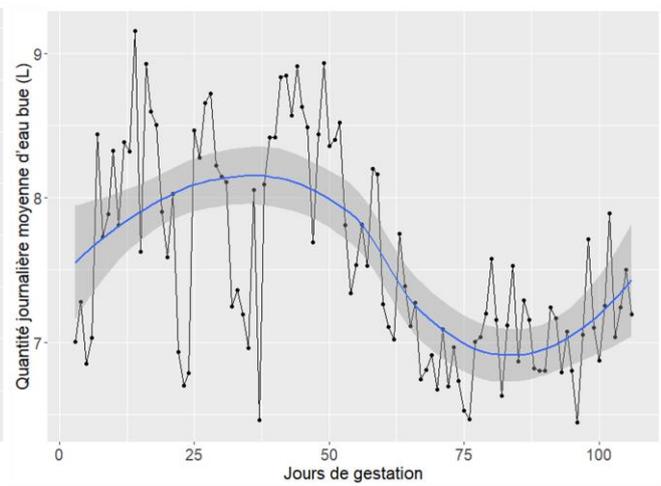
La durée moyenne des buvées d'une truie en situation basale est de 8 min/jour. Il existe cependant une forte variabilité intra-individuelle (*Figure 4*).

La *Figure 3* montre une importante variabilité inter-truies d'une journée à l'autre concernant les habitudes de consommation d'eau et d'activité. Cette grande variabilité ne permet pas d'établir un système d'alerte à partir d'un seuil fixe qui serait calculé à partir de « l'individu moyen » de la bande.

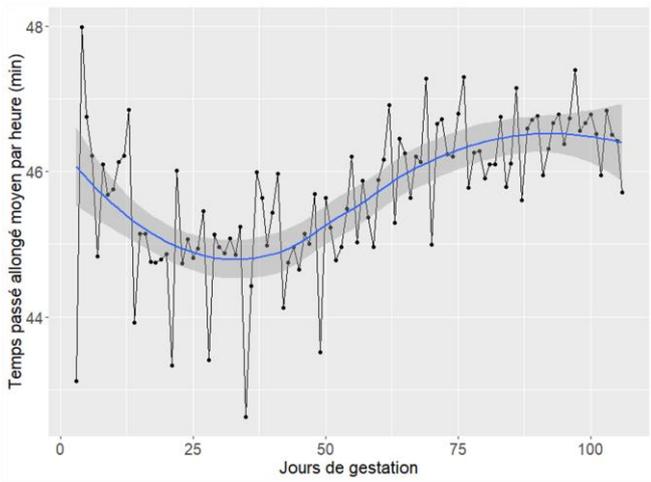
Il faudrait pouvoir individualiser ces seuils en situation basale sans évènement, afin d'identifier les écarts consécutifs à des évènements inhabituels.



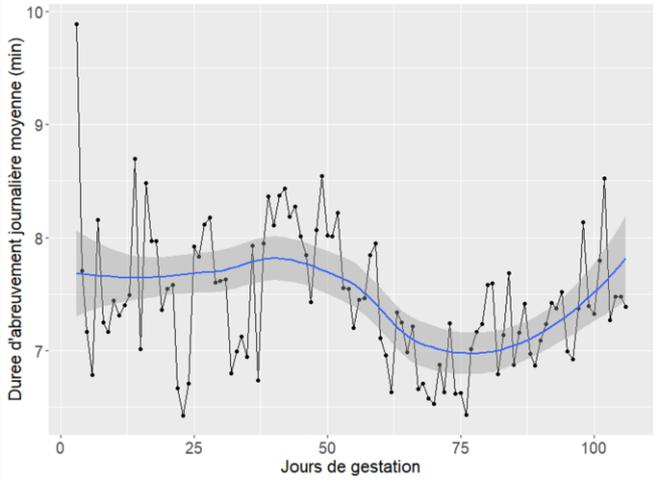
3.a



3.b



3.c



3.d

Figure 3 : Moyenne par truie et par jour de gestation en situation basale pour les (3.a) Nombre de buvées, (3.b) Quantité d'eau bue (L), (3.c) Temps passé allongé (min), (3.d) Durée d'abreuvement (min)
 La courbe bleue représente la courbe de tendance obtenue par la moyenne mobile
 L'aire grisée représente l'intervalle de confiance à 95%

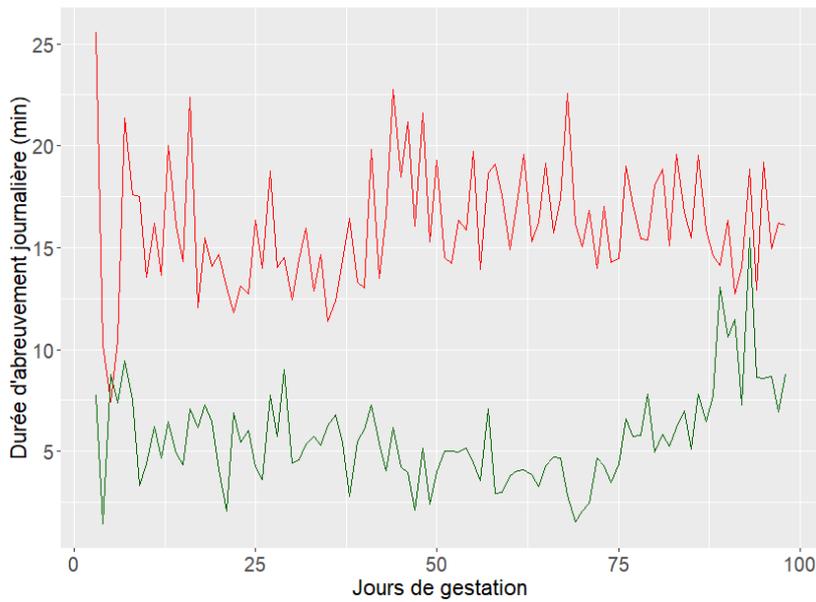


Figure 4: Variations intra-individuelles de la durée d'abreuvement journalier pour 2 truies (en rouge ou vert) de la bande 17

3.2. Evènements perturbateurs

Des évènements thermiques (chaud et froid) ou alimentaires (Alim1 et Alim2) ont été induits sur 2x2 bandes de truies. Des boxplots réalisés pour chacune des 6 variables quantitatives et selon chaque évènement, montrent l'étendue des valeurs (voir Figure 5). Ainsi, on constate que la quantité d'eau bue par jour et par truie peut varier de moins d'un litre à plus de 45 litres/jour. La moyenne se situe aux alentours de 8 litres d'eau/jour/truie. Quand il fait chaud, les truies boivent en moyenne plus d'eau. Quand il fait froid ou en période de perturbation alimentaire, les truies restent en moyenne allongées moins longtemps par rapport à la situation basale qui est d'environ 45 min/heure.

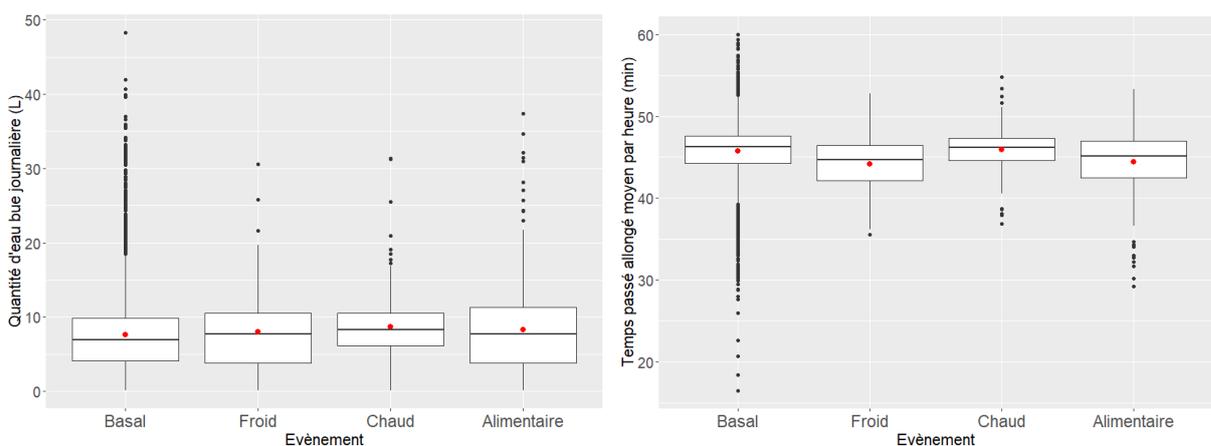


Figure 5 : Boxplots de la quantité d'eau bue (L) et du temps passé allongé (min) par jour selon l'Evènement

3.2.1. Evènement thermique Chaud

Les truies de la bande 6 soumises à un évènement thermique chaud, ont consommé en moyenne 9.79 ± 5.80 litres/jour d'eau alors qu'à la même période de gestation (entre J35 et J38), en situation basale, elles auraient consommé 8.08 ± 0.11 litres/jour en moyenne ($pval = 0.018$). Ces résultats sont en accord avec la littérature qui soutient qu'une hausse de la température extérieure entraîne une augmentation de la consommation d'eau (Bird et al., 2001).

Les truies de la bande 6 ont également passé plus de temps à boire (9.63 ± 4.99 min/jour en moyenne) lors de l'évènement chaud, au lieu de 7.69 ± 0.005 min en situation basale sur la même période de gestation ($pval = 0.007$). Elles se sont rendues plus souvent à l'abreuvoir. Elles ont aussi changé plus souvent de position en période de forte chaleur par rapport à la situation basale à une température ambiante d'environ 20°C durant la même période de gestation : 1.80 ± 0.49 fois/heure vs. 1.53 ± 0.002 fois/heure en moyenne respectivement ($pval = 0.006$).

Les truies de la bande 5 ont passé moins de temps debout 9.86 ± 2.84 min/heure en moyenne durant l'évènement chaud au lieu de 11.15 ± 0.03 min/heure à la même période de gestation en situation basale. Elles ont adopté un comportement différent face à l'évènement thermique chaud, en restant plus couchées ventralement vs. latéralement afin de davantage bénéficier de la fraîcheur du sol (Abarnou et al., 2022).

On constate un effet bande qui peut s'expliquer par le fait que les salles de gestation étaient différentes selon la bande 5 ou 6. En effet, la salle dans laquelle se trouvaient les truies de la bande 6 est mieux isolée thermiquement, ce qui expliquerait pourquoi elles ont davantage ressenti et réagi aux évènements thermiques. D'autre part, les perturbations thermiques sont intervenues plus tôt dans l'avancée de la gestation (J21 pour la bande 6 vs. J42 pour la bande 5) quand les truies sont plus mobiles (*Figure 1*).

3.2.2. Evènement thermique Froid

On constate à nouveau un effet bande. En effet, les truies de la bande 6 soumises à un évènement thermique froid restent en moyenne allongées moins longtemps (43.51 ± 2.85 min/heure) par rapport à une situation basale (45.05 ± 0.028 min/heure) au même stade de gestation. Nous avons ainsi pu mettre en évidence un changement significatif dans la posture des truies en situation d'évènement thermique froid ($pval = 0.0058$).

Le fait qu'elles soient plus debout quand il fait froid est en accord avec la littérature qui montre qu'à une température ambiante froide, les truies passent significativement plus de temps à manger et à manipuler la paille qu'à une température moyenne (Hayne et al., 2000).

3.2.3. Evènement Alimentaire

L'évènement de perturbation alimentaire a entraîné une augmentation significative de l'activité des truies dans les 2 bandes. En effet, par rapport à une situation basale, elles sont restées significativement plus longtemps debout 14.16 ± 4 min/heure vs. 11.57 ± 0.007 min/heure ($pval = 0.015$).

Par ailleurs, les truies de la bande 16 ont bu significativement plus d'eau ($pval = 0.036$).

On peut supposer que l'évènement alimentaire a pu mettre en exergue l'effet de la hiérarchie dans la bande de truies. En effet, dans les élevages porcins, la constitution de la bande de truies en début de gestation, entraîne l'établissement d'une hiérarchie (Fraser et Rushen, 1987) avec un risque plus élevé de blessures à cause des agressions (Puppe et al., 2008). Ce phénomène de

hiérarchie dans le groupe peut conduire à une difficulté d'accès aux abreuvoirs (Andersen et al., 2008) et aux DAC.

3.2.4. Evènement Malade

Durant les expérimentations, 3 truies ont été malades. Pour une de ces truies, on peut constater sur la *Figure 6* une forte hausse du nombre des buvées dès J48. Cette truie a été malade entre J49 et J51 (boiterie constatée et administration d'antibiotiques). Elle a consommé plus d'eau en passant d'une consommation de 3 litres/jour avant l'infection à plus de 6 litres/jour juste avant l'administration des médicaments. Ce résultat est en accord avec la littérature, où il est mentionné qu'en élevage porcin la consommation d'eau augmente dans le cas de certaines maladies métaboliques ainsi qu'en réponse à la déshydratation résultant de la diarrhée (Madsen et Kristensen, 2005; Seddon et al., 2011). Il a aussi été démontré que la durée et la fréquence des buvées ainsi que le temps passé allongé prédisent les stades précoces des infections bactériennes (Ivos et al., 1981).

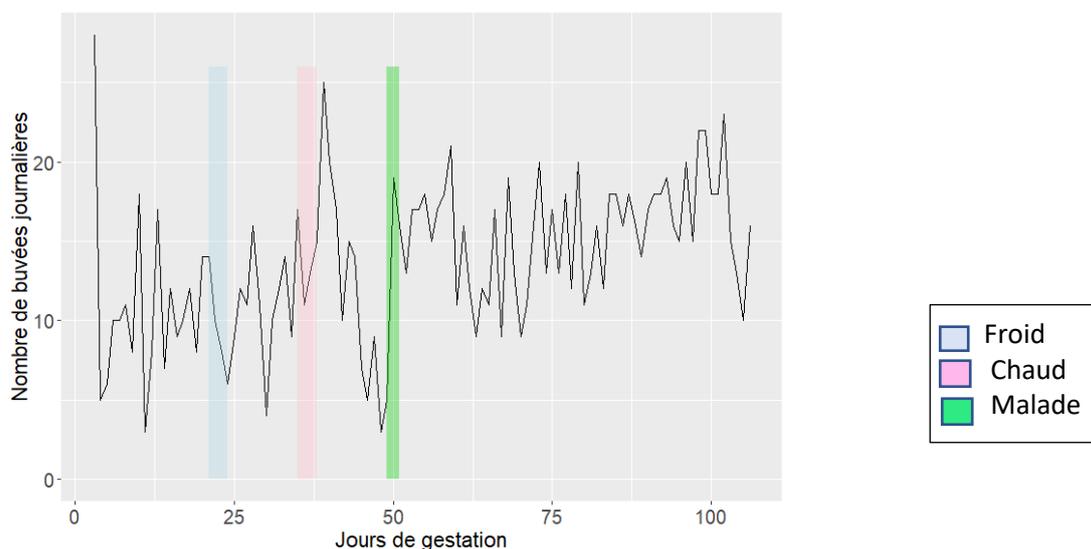


Figure 6 : Fréquence d'abreuvement d'une truie de la bande 6 malade pendant la gestation

L'ensemble de ces résultats démontre l'importance de la prise en compte de plusieurs variables liées à la fois au comportement d'abreuvement et à l'activité des truies dans le cadre de la surveillance active des élevages (voir *Tableau 3*). De cette façon, les abreuvoirs connectés, les accéléromètres et les capteurs associés aux DAC devraient permettre de détecter plus efficacement les évènements perturbateurs survenant en élevage de truies (Marcon et al. 2020). Ces premiers résultats indiquent que l'évènement alimentaire est mieux détecté par les variables « debout » ou « allongé » et « quantité d'eau bue ». L'évènement chaud est plus facile à détecter en suivant les variables « durée d'abreuvement », « quantité d'eau bue » et « nombre de changements de position ». L'évènement froid occasionne chez les truies une durée « allongée » plus courte ou « debout » plus longue.

Evènement Variable	Froid	Chaud	Alimentaire
Debout	+ 1.23 (± 0.29) *	- 0.07 (± 0.29)	+ 1.22 (± 0.34) *
Allongée	- 2.08 (± 0.37) *	+ 0.095 (± 0.37)	- 1.36 (± 0.32) *
Change position	- 0.05 (± 0.04)	+ 0.13 (± 0.04) *	- 6.88e-03($\pm 2.68e-02$)
Durée	- 0.80 (± 0.52)	+ 1.42 (± 0.52) *	+ 0.15 (± 0.29)
Quantité	- 0.02(± 0.60)	+ 1.38 (± 0.60) *	+ 0.80 (± 0.31) *
Nb buvées	- 0.52 (± 0.80)	+ 0.15 (± 0.80)	+ 0.23 (± 0.38)

* Pvalue ≤ 0.05

Tableau 3 : Profils d'évolution des différences entre la moyenne de chaque variable par rapport à la moyenne en situation basale pendant l'évènement

Afin de détecter les potentielles variations dans les courbes de suivi temporel individuel des truies pour chacune des variables mesurées, nous avons choisi de tester 3 méthodes différentes.

3.3. Méthode de lissage différentiel

Rappelons que par truie et par variable étudiée, une pénalité forte de 100 est appliquée sur la dérivée des fonctions. Elle permet d'obtenir une courbe plus lissée limitant l'amplitude des variations. Plus la pénalité est forte et plus la courbe est lissée Cette courbe très lissée est appelée la "**courbe théorique**".

D'autre part, une pénalité faible de 10 est appliquée sur la dérivée des fonctions afin de mettre en évidence les variations les plus importantes (par rapport à la courbe très lissée). Cette courbe moins lissée est nommée la "**courbe réalisée**". Sur la *Figure 7*, les différences entre les courbes théoriques (*en pointillé violette*) et les courbes réalisées (*continue noire*) permettent de détecter les perturbations imputables à des évènements alimentaires ou thermiques.

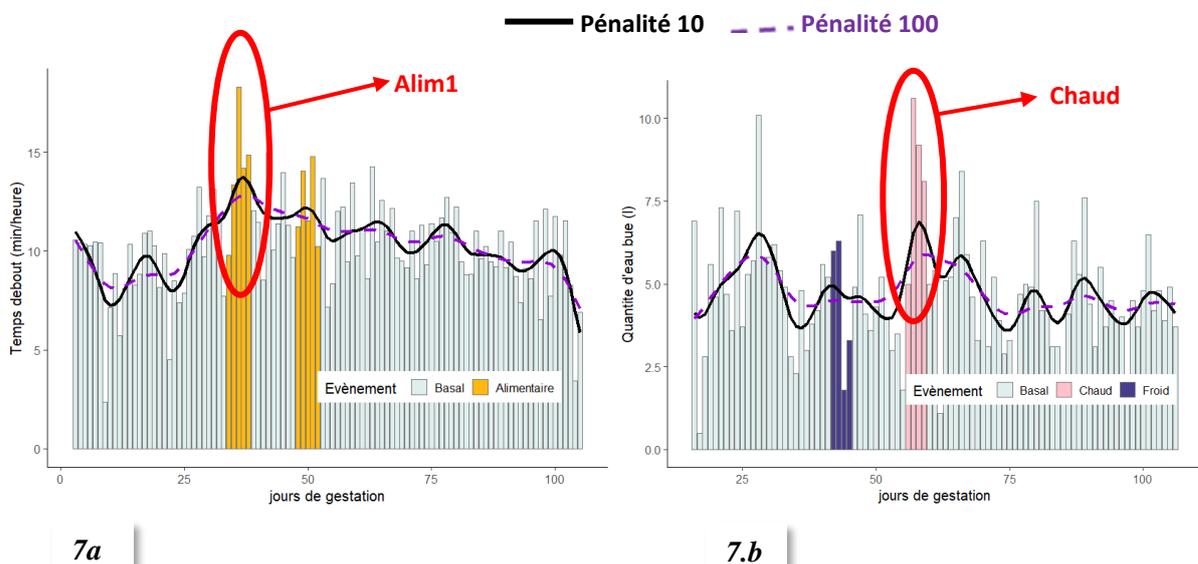


Figure 7 : (7.a) Temps passé debout pour une truie de la bande 16 pendant la gestation (7.b) Quantité d'eau bue pour une truie de la bande 5 pendant la gestation

Les résultats obtenus pour chacune des 6 variables d'abreuvement et d'activité sont présentés dans le *Tableau 4*. On peut remarquer que la prise en compte de " l'aire" entre les 2 courbes de lissage donne presque toujours de meilleurs résultats que le "max_diff "ou la "moy_diff "en ce qui concerne les valeurs de précision, de sensibilité, de spécificité et de VPP.

La précision est plus élevée avec le quantile 0.95 qu'avec le quantile 0.85, en effet, le seuil de positivité fixé à 5% des plus fortes valeurs absolues des aires entre les courbes, max_diff ou moy_diff est plus restrictif qu'à 15% et permet une meilleure sélection des perturbations dues aux évènements par rapport aux non évènements. Les variables « temps allongé », « fréquence des buvées », « durée des buvées » et « quantité des buvées » sont par ordre décroissant en terme de précision les plus importantes pour détecter les perturbations liées aux évènements.

Les **sensibilités maximums** obtenues sont de 40.37 % pour le **quantile 0.85** et de 18.02 % pour le quantile 0.95. Les **spécificités maximums** obtenues sont de 92.44 % pour le **quantile 0.95** et de 73.16 % pour le quantile 0.85. Par exemple, pour la variable « temps allongé », avec le quantile 0.95, la spécificité est de 92.44 % tandis que la sensibilité est de 18.02 %. Avec le quantile 0.85, toujours pour la variable « temps allongé », la sensibilité est de 40.28 % tandis que la spécificité est de 72.90 %.

Par ailleurs, les valeurs prédictives positives (VPP) se situent entre 27.61 % et 44% (aire). Cela signifie que quand la méthode détecte une perturbation, il y a réellement un évènement dans un tiers à presque la moitié des cas.

quantile 0.95 (a)

		Fréquence buvées	Durée buvées	Temps debout	Changement position	Temps allongé	Quantité buvées
Précision (%)	aire	72.11	71.48	68.84	70.36	74.01	71.19
	max-diff	69.23	70.99	68.54	68.08	70.86	70.62
	moy-diff	68.94	69.35	66.46	67.97	70.34	68.97
Sensibilité (%)	aire	15.09	15.38	10.82	13.67	18.02	13.85
	max-diff	9.43	14.28	9.70	10.15	11.66	13.10
	moy-diff	9.05	11.72	6.34	8.59	11.66	10.86
Spécificité (%)	aire	91.61	91.68	89.77	90.83	92.44	91.05
	max-diff	89.67	91.42	89.77	88.99	90.34	90.62
	moy-diff	89.41	90.10	88.15	89.42	89.65	89.10
VPP (%)	aire	38.09	40	27.61	35	43.96	34.90
	max-diff	23.80	37.5	25.49	25	28.44	32.71
	moy-diff	22.64	29.90	16.19	22.68	27.04	25.66
Prop.Evèn. (%)		25.48	26.47	26.50	26.52	24.76	25.72

quantile 0.85 (b)

		Fréquence buvées	Durée buvées	temps debout	changement position	Temps allongé	Quantité buvées
Précision (%)	aire	64.80	63.72	62.21	63.21	64.82	62.90
	max-diff	62.11	61.97	61.12	60.82	70.51	63.96
	moy-diff	61.25	60.91	59.44	60.62	61.15	60.11
Sensibilité (%)	aire	40.37	38.46	35.07	37.89	40.28	36.70
	max-diff	34.71	34.79	32.83	32.81	33.21	38.20
	moy-diff	32.83	32.23	29.47	32.03	32.15	30.71
Spécificité (%)	aire	73.16	72.82	72	72.35	72.90	71.98
	max-diff	71.48	71.76	71.33	70.94	71.27	72.89
	moy-diff	70.96	71.24	70.25	70.94	70.69	70.29
VPP (%)	aire	33.96	33.76	31.12	33.10	32.85	31.21
	max-diff	29.39	30.74	29.23	28.96	27.56	32.79
	moy-diff	27.88	28.75	26.33	28.47	26.53	26.36

Tableau 4 : Mesure de la performance des tests par variable tous Evènements confondus pour le quantile 0.95 (a) et le quantile 0.85 (b) (méthode du lissage différentiel), toutes bandes confondues Les valeurs importantes sont surlignées en jaune

Les résultats des tests par bande, toutes variables confondues, selon l'évènement (Tableau 5) montrent que le 1^{er} évènement est mieux détecté que le 2nd évènement lorsqu'on regarde les sensibilités et les VPP obtenues. Comme vu précédemment, la précision et la spécificité sont meilleures avec le quantile 0.95 et la sensibilité est plus forte avec le quantile 0.85. Les VPP se situent entre 14% et 35%. La méthode permet de mieux détecter les évènements thermiques en terme de précision et de spécificité. Par contre, en terme de sensibilité, la méthode détecte mieux les évènements perturbateurs alimentaires.

quantile 0.95 (a)

		Bande 5	Bande 6			Bande 16	Bande 17
Froid	Précision (%)	81.92	82.07	Alim1	Précision (%)	77.72	78.17
	Sensibilité (%)	18.81	11.11		Sensibilité (%)	18.89	19.67
	Spécificité (%)	93.55	94.02		Spécificité (%)	88	89.13
	VPP (%)	35	23.85		VPP (%)	21.57	25.35
Chaud	Précision (%)	81.61	82.77	Alim2	Précision (%)	77.96	77.57
	Sensibilité (%)	11.11	9.38		Sensibilité (%)	13.02	14.52
	Spécificité (%)	93.55	94.02		Spécificité (%)	88	89.13
	VPP (%)	22.61	19.41		VPP (%)	14.36	19.69

quantile 0.85 (b)

		Bande 5	Bande 6			Bande 16	Bande 17
Froid	Précision (%)	72.05	70.54	Alim1	Précision (%)	64.15	64.02
	Sensibilité (%)	39.24	31.62		Sensibilité (%)	43.31	51.36
	Spécificité (%)	78.09	77.1		Spécificité (%)	67.79	66.39
	VPP (%)	24.83	18.87		VPP (%)	19.02	22.27
Chaud	Précision (%)	71.77	71.91	Alim2	Précision (%)	63.24	61.81
	Sensibilité (%)	34.5	38.02		Sensibilité (%)	33.85	36.87
	Spécificité (%)	78.09	77.1		Spécificité (%)	67.79	66.39
	VPP (%)	21.07	20.3		VPP (%)	13.97	16.75

Tableau 5 : Mesure de la performance des tests par Bande selon l'Évènement pour le quantile 0.95(a) et le quantile 0.85 (b) (méthode du lissage différentiel)
Les valeurs importantes sont surlignées en jaune

En conclusion, on peut dire qu'il est possible par la méthode du lissage différentiel de détecter des événements perturbateurs induits via les modifications des dynamiques d'abreuvement et d'activité physique des truies. Cependant, les sensibilités obtenues restent faibles (max 51% avec le quantile 0.85) pour permettre une utilisation suffisamment fiable en élevage. Les spécificités obtenues sont bonnes avec le quantile 0.95 mais au détriment de la sensibilité. Ces résultats montrent qu'il est important de trouver un compromis entre la sensibilité et la spécificité par le biais du choix du quantile (seuil de positivité de la détection d'une perturbation).

3.4. Tests non paramétriques de détection de ruptures de type *CUSUM*

Les résultats (*Tableau 6*) montrent que la méthode a permis la détection d'évènements grâce à la mise en évidence de points de rupture de corrélation entre les couples de variables pendant les événements perturbateurs. Nous pouvons remarquer que comme avec la méthode du lissage différentiel, le 1^{er} évènement est mieux détecté que le 2nd. Ceci pourrait être attribué au fait que le 2nd évènement intervient seulement 10 jours après le 1^{er} alors que les truies n'ont pas totalement récupéré. Il est alors plus difficile d'identifier clairement le début du 2nd évènement.

La sensibilité la plus forte est de 42.85% et la VPP la plus forte est de 66.66 % pour la Bande 17 suite à l'évènement perturbateur Alim1. On peut remarquer que la méthode a permis une meilleure détection des évènements alimentaires que des évènements thermiques et plus particulièrement du 1^{er} évènement alimentaire.

D'autre part, les variables « quantité des buvées », « temps allongé » et « temps debout » sont presque toujours retrouvées dans les couples de variables ayant permis de détecter un ECP pendant un évènement alimentaire (VP). Ceci est en accord avec les résultats des tests d'analyse de variance (*Tableau 3*) qui montrent une différence significative entre les moyennes de chacune de ces 3 variables par rapport à la moyenne en situation basale pendant l'évènement alimentaire.

Nous pouvons donc confirmer notre hypothèse initiale qui supposait que la liaison entre certaines variables change en période d'évènement et plus particulièrement chez certaines truies.

		Bande 5	Bande 6			Bande 16	Bande 17
Froid	Sensibilité (%)	18.38	21.05	Alim1	Sensibilité (%)	41.05	42.85
	VPP (%)	20	44.44		VPP (%)	53.33	66.66
Chaud	Sensibilité (%)	15.18	18.18	Alim2	Sensibilité (%)	33.33	30
	VPP (%)	16.66	13.33		VPP (%)	35.71	30

Tableau 6 : Mesure de la performance des tests par Bande selon l'Evènement (méthode de détection de ruptures)

Les valeurs importantes sont surlignées en jaune

3.5. Random Forest

L'analyse des données par Random Forest sur chaque bande de truies a permis de montrer que les variables les plus importantes par ordre décroissant sont pour la bande 6 : « changement de position », « durée d'abreuvement », « temps allongé » et « quantité d'eau bue » et pour la bande 16 : « temps allongé », « temps debout » et « quantité d'eau bue ».

Les résultats sur le *Tableau 7* montrent que la sensibilité maximale atteint 53.33 %, avec une précision de 92.91 %, une spécificité de 99.45% et une VPP de 75% pour la bande 6 lorsque les truies ont été soumises à l'évènement thermique froid.

On constate un effet bande entre les bandes 5 et 6 surtout lors de l'évènement perturbateur froid en terme de sensibilité. Ceci peut s'expliquer par le fait que les truies de la bande 6 qui étaient logées dans une salle mieux isolée thermiquement, ont davantage ressenti les évènements thermiques induits.

On remarque également un effet bande entre les bandes 16 et 17 à la suite des perturbations alimentaires. Effectivement, la précision et la sensibilité sont plus élevées pour la bande 16.

Le fait qu'il y avait plus de truies multipares (ayant déjà eu au moins une portée) dans la bande 16 que dans la bande 17 (83.34% vs. 60%), pourrait expliquer des différences d'activité entre ces 2 bandes. En effet, il a été rapporté que les truies multipares ont un niveau d'activité plus élevé (60% plus actives) que les truies primipares (Cariolet et Dantzer, 1984).

Enfin, on remarque là encore que les premiers évènements induits sont mieux détectés que les seconds. Le 2nd évènement intervenant seulement 10 jours après le 1^{er}, on peut supposer que les truies n'aient pas eu le temps de retrouver un niveau basal selon la variable considérée. Ceci est d'autant plus vrai pour le 2^{ème} évènement alimentaire qui, identique au précédent, a permis aux truies une certaine « habitude ».

		Bande 5	Bande 6			Bande 16	Bande 17
	Précision (%)	94.42	92.91		Précision (%)	92.17	85.26
Froid	Sensibilité (%)	33.33	53.33	Alim1	Sensibilité (%)	40	9.09
	Spécificité (%)	99.61	99.45		Spécificité (%)	98.48	100
	VPP (%)	75	75		VPP (%)	54.54	100
Chaud	Sensibilité (%)	25	26.66	Alim2	Sensibilité (%)	20	0
	Spécificité (%)	99.22	100		Spécificité (%)	99.69	98.85
	VPP (%)	60	100		VPP (%)	80	0

Tableau 7 : Mesure de la performance des tests par Bande selon l'Évènement (Random Forest)
Les valeurs importantes sont surlignées en jaune

En conclusion, on peut dire d'après l'ensemble des indicateurs de performance que la méthode du Random Forest est meilleure pour détecter les événements thermiques que les 2 autres méthodes et équivalente à la méthode de détection des ruptures de corrélation pour détecter les événements alimentaires (bande 16) en terme de sensibilité et de VPP.

4. Conclusion et perspectives

L'analyse des données issues des abreuvoirs connectés et des accéléromètres nous a permis de confirmer que les variables liées aux comportements d'abreuvement et d'activité physique des truies représentent de bons indicateurs pour mettre en évidence la survenue de problèmes dans les élevages porcins. Dans le cadre de notre étude, les principaux marqueurs capables de détecter les événements thermiques et alimentaires sont le temps passé debout (ou allongé), le nombre de changements de position et la quantité d'eau bue. Le profil d'évolution de ces variables par rapport à l'état basal est caractéristique de l'évènement.

Cependant, chez certaines truies, les événements induits n'ont pas entraîné suffisamment de modifications détectables au niveau des courbes temporelles pour certaines variables par rapport à l'état basal assez fluctuant.

La détection des événements perturbateurs thermiques est meilleure avec la méthode des forêts aléatoires. Les sensibilités restent néanmoins assez faibles (53% max pour l'évènement thermique froid). Les événements alimentaires sont aussi bien identifiés par la méthode de détection des ruptures de corrélation que par la méthode du Random Forest.

La performance des tests pourrait certainement être améliorée en augmentant la taille du jeu de données. Effectivement, nous avons réalisé les analyses à partir de 63 truies gestantes et il serait intéressant de disposer des enregistrements d'un plus grand nombre d'animaux pour améliorer la robustesse des tests.

Par ailleurs, la méthode du lissage différentiel pourrait être optimisée en choisissant un quantile seuil qui maximise le compromis entre la sensibilité et la spécificité.

D'autre part, le design expérimental pourrait être amélioré en espaçant les 2 évènements successifs de plus de 10 jours. L'idéal serait de ne faire subir aux truies qu'un seul évènement perturbateur sur une période de gestation afin d'éviter les biais dus à l'effet rémanent du 1^{er} évènement sur le 2nd.

De plus, étant donné la grande variabilité entre les individus et afin de permettre à terme une détection précoce des signaux avant-coureurs de problèmes de santé, de dysfonctionnements matériels ou de mal-être animal, il faudrait disposer des données individuelles, sur une période de gestation complète, en situation basale, sans évènement perturbateur. En effet, les habitudes individuelles de consommation d'eau et d'activité physique seraient mieux prises en compte et permettraient d'établir des courbes témoins de référence propres à chaque animal. Il serait également intéressant de tester d'autres méthodes de détection d'évènements comme les tests globaux de variance fonctionnelle.

La mise en place de ces améliorations rendrait ainsi plus fiable la gestion des alertes en cas d'anomalies survenant dans les élevages.

Bibliographie

Abarnou, J., Durand, M., Dourmad, J-Y., Gaillard, C., 2022. Effects of induced thermal conditions on gestating sows' behaviors and energy requirements (*soumis à Journal of Animal Science*)

Andersen, H.M.-L., Dybkjær, L., Herskin, M.S., 2014. Growing pigs' drinking behaviour: number of visits, duration, water intake and diurnal variation. *Animal* 8 (11), 1881-1888. DOI 10.1017/S175173111400192X

Andersen, H.M.-L., Jørgensen, E., Dybkjær, L., Jørgensen, B., 2008. The ear skin temperature as an indicator of the thermal comfort of pigs. *Applied Animal Behaviour Science* 113 (1-3), 43-56. DOI 10.1016/j.applanim.2007.11.003

Bird, N., Crabtree, H., Scholfield, C.P., 2001. Engineering technologies enable real time information monitoring in pig production. British Society of Animal Science Occasional Publication. ed 105-112

Cariolet, R., Dantzer, R., 1985. Moving activity of pregnant tethered sows- determination of some factors of variations. *Annales de zootechnie.* vol 34

Cerneau, P., Meunier Salaun, M.C., Lauden, P., Godfrin, K., 1997. Effects of housing and feeding practices on the behaviour and reproductive performance of pregnant sows. Agris FAO

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. DOI 10.2527/jas.2010-3753

Cornou, C., Lundbye- Kristensen, S., 2012. Modeling of sows diurnal activity pattern and detection of parturition using acceleration measurements. *Computers and Electronics in Agriculture* (80), 97-104

Cornou, C., Ringgaard Kristensen, A., 2014. Monitoring individual activity before, during and after parturition using sensors for sows with and without straw amendment. *Livestock Science* (168), 139-148

Courboulay, V., Marcon, M., 2017. Suivi Individuel Des Truies Gestantes Pour Améliorer Leur Bien-Être et Leurs Performances. *IFIP, Bilan d'activité*

Fraser, D., Rushen, J., 1987. Aggressive Behavior. *Veterinary Clinics of North America: Food Animal Practice* 3 (2), 285-305. DOI 10.1016/S0749-0720(15)31153-1

Gaillard, C., Friggens, N. C., Taghipoor, M., Weisbjerg, M. R., Lehmann, J. O., Sehested, J., 2015. Effects of an individual weight-adjusted feeding strategy in early lactation on milk production of Holstein cows during extended lactation. *J. Dairy Science.* 99:2221–2236

Gaillard, C., Dourmad, J-Y., 2022. Application of a precision feeding strategy for gestating sows. *Animal Feed Science and Technology.* 287
<https://doi.org/10.1016/j.anifeedsci.2022.115280>

Hayne, S. M., Tennessen, T., Anderson, D.M., 2000. The responses of growing pigs exposed to cold with varying amounts of straw bedding. *The Canadian veterinary journal* 80(4):539-546

Junge, M., Herd, D., Jezierny, D., Gallmann, E., Jungbluth, T., 2012. Indicators for monitoring behavior and health of group housed pregnant sows. *Landtechnik* 67, 326-331

Klopfenstein, C., Bigras-Poulin, M., Martineau, G.-P., 1996. La truie potomane, une réalité physiologique. *Journées de Recherche Porcine*. ed 319-324

Kojadinovic, Ivan, and Maintainer Ivan Kojadinovic. "Package 'npcp'." *Econometric Reviews* 23.1 (2020): 53-70

Kojadinovic, I., Quessy, J. F., Rohmer, T., 2016. Testing the constancy of Spearman's rho in multivariate time series. *Annals of the Institute of Statistical Mathematics*, 68(5), 929-954

Liam, A., Wiener, M., Package 'randomForest'. 2002 "Classification and Regression by randomForest." *R News*, 2(3), 18-22. <https://CRAN.R-project.org/doc/Rnews/>

Madsen, T.N., Kristensen, A.R., 2005. A model for monitoring the condition of young pigs by their drinking behaviour. *Computers and Electronics in Agriculture* 48 (2), 138-154. DOI 10.1016/j.compag.2005.02.014

Marcon, M., Quiniou, N., Courboulay, C., Rousselière, Y., Melot, G., Meunier-Salaün, M.C., Labuissière, E., Ramonet, Y., Cherel, P., Le Mer, M., 2020. Améliorer les performances et le bien être des truies gravides par la mobilisation de nouvelles technologies pour une alimentation de précision et la détection de signaux comportementaux. *Innovations Agronomiques* 79, 245-256

Martineau, G.-P., Morvan, H., 2010. Maladies d'élevage des porcs, *France Agricole*

Mateu Chaler, S., 2020. Importance de la consommation d'eau chez les truies pendant la lactation
https://www.3trois3.com/articles/importance-de-la-consommation-deau-des-truies-pendant-la-lactation_14600

Puppe, B., Langbein, J., Bauer, J., Hoy, S., 2008. A comparative view on social hierarchy formation at different stages of pig production using sociometric measures. *Livestock Science* 113 (2-3), 155-162. DOI 10.1016/j.livsci.2007.03.004

R Core Team (2021). R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. URL <https://www.R-project.org>

Ramsay, J. O., Wickham Hadley, Ramsay, Maintainer JO, *et al.* Package 'fda'. 2022

Ramsay, J.O., Graves, S., Hooker, G., 2020. *fda: Functional Data Analysis*

Rousselière, Y., Hémonic, A., Marcon, M., 2017. Monitoring of the individual drinking behavior of healthy weaned piglets and pregnant sows

Seddon, Y., Farrow, M., Guy, J., Edwards, S., 2011. Can monitoring water consumption at pen level detect changes in health and welfare in small groups of pigs? *Proceedings of the 5th International Conference on the Assessment of Animal Welfare at Farm and Group Level* (13)

	Diplôme : Master 2 Spécialité : Spécialisation / option : Science des données Enseignant référent : Causeur David
Auteur(s) : Hamelin Gaëlle	Organisme d'accueil : INRAE Adresse : UMR Pegase Le Clos 35590 Saint-Gilles
Nb pages : 20 Annexe(s) :	
Année de soutenance : 2022	Maîtres de stage : Charlotte Gaillard Tom Rohmer
Amélioration d'un outil de détection d'évènements (techniques ou sanitaires) à l'échelle de plusieurs bandes de truies	
Improvement of an event detection tool (technical or sanitary) at the scale of several groups of sows	
<p>L'agriculture numérique s'est considérablement développée ces dernières années et plus particulièrement dans l'élevage porcin. Il est actuellement possible de suivre les animaux individuellement pour détecter l'apparition de problèmes techniques ou sanitaires. En effet, le suivi individuel des animaux d'élevage permet une remontée d'informations et représente un outil d'aide à la décision en complément des observations qui peuvent être faites. L'objectif de cette étude a été de tester et de comparer 3 méthodes de détections d'évènements (techniques ou sanitaires) à partir des données temporelles issues de capteurs d'abreuvement et d'activité physique placés à l'oreille des truies gestantes. Le but à long terme est de réussir à développer des outils automatisés les plus fiables possibles permettant de déceler rapidement les situations au cours desquelles des variations du comportement d'abreuvement et de l'activité physique peuvent être liées à un événement sanitaire ou technique.</p> <p>Les données ont été collectées à partir de 63 truies gestantes à l'UE3P de l'INRAE de Saint Gilles. Quatre bandes de truies ont subi soit 2 évènements thermiques consécutifs, soit 2 évènements alimentaires rapprochés. Dans cette étude, les 3 méthodes de détection d'évènements testées et comparées sont une méthode de lissage différentiel, une méthode de détection des ruptures de corrélations entre les différents couples de variables ou la méthode des forêts aléatoires. Les premiers résultats montrent que lors d'une augmentation de température, les truies changent plus souvent de position et ingèrent plus d'eau, plus longtemps. En cas de compétition alimentaire, elles passent moins de temps allongées qu'en situation basale et boivent plus d'eau. Les meilleurs résultats sont obtenus avec les forêts aléatoires.</p>	
<p>Digital agriculture has grown considerably in recent years, especially in pig farming. It is currently possible to track animals individually to detect the occurrence of technical problems or health events. In fact, the individual tracking of farm animals allows information to be gathered and can serve as a decision tool. The objective of this study was to test and compare 3 methods of detecting events (technical or sanitary) from temporal data from water troughs and physical activity sensors placed in the ear of pregnant sows. The long-term goal is to successfully develop the most reliable automated tool possible to quickly identify variations in drinking behavior and/or physical activity that could indicate a health or technical event.</p> <p>The data were collected from 63 pregnant sows at the INRAE UE3P of Saint Gilles. Four groups of sows experienced either 2 consecutive thermal disturbances or 2 close dietary disturbances. In this study, the 3 methods of detection of events tested and compared are a method of differential smoothing, a method of detection of ruptures of correlations between the different pairs of variables or random forests. The first results show that when temperature rises, sows change position more often and ingest more water longer. In case of a feed competition, they spend less time lying down than in a basal situation and drink more water. The best results are obtained with random forests.</p>	
Mots-clés : élevage de précision, truies gravides, accéléromètres, abreuvoirs connectés, capteurs, évènements techniques et sanitaires, traitement dynamique des données temporelles	
Key Words: precision breeding, gravid sows, accelerometers, connected water troughs, sensors, technical and sanitary events, dynamic processing of temporal data	