

Machine Learning applied to behaviour monitoring to detect diseases, reproductive events and disturbances in dairy cows

Isabelle Veissier, Romain Lardy, Marie-Madeleine Mialon, Q. Ruin, V. Antoine, J. Koko

▶ **To cite this version:**

Isabelle Veissier, Romain Lardy, Marie-Madeleine Mialon, Q. Ruin, V. Antoine, et al.. Machine Learning applied to behaviour monitoring to detect diseases, reproductive events and disturbances in dairy cows. 26. Rencontres autour des Recherches sur les Ruminants (3R 2022), Institut de l'Elevage; INRAE, Dec 2022, Paris, France. pp.562-564. hal-03943072

HAL Id: hal-03943072
https://hal.inrae.fr/hal-03943072

Submitted on 17 Jan 2023

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Machine Learning appliqué au suivi du comportement pour identifier maladies, états reproductifs et perturbations des vaches laitières

VEISSIER I. (1), LARDY R. (1), MIALON M.M. (1), RUIN Q. (1), ANTOINE V. (2), KOKO J. (2)
(1) Université Clermont Auvergne, INRAE, VetAgro Sup, UMR Herbivores, 63122 Saint-Genès-Champanelle, France
(2) Université Clermont Auvergne, CNRS, UMR 6158 LIMOS, Campus des Cézeaux – BP 10125, 63173 Aubiière cedex, France

RESUME
Le comportement d'un individu est très sensible à toute perturbation extérieure ou modification de l'état interne. Nous avons cherché à prédire et à classer un large éventail d'états de santé, de stress et d'états physiologiques à partir de l'activité quotidienne de vaches laitières. Nous disposions de six jeux de données (au total 122 000 vaches*jours). Les soigneurs notaient l'état des vaches : maladie, œstrus, vêlage, perturbations (manipulations, mélange). Les vaches étaient équipées de capteurs permettant d'estimer la durée des activités - manger, se reposer, dans les couloirs – à partir desquelles était calculé le niveau d'activité. Nous avons modélisé le rythme d'activité par des transformées de Fourier (Harmoniques 0 et 1). Si l'écart entre les modèles obtenus au cours de 2 séries de 24 h est supérieur à un certain seuil, il y a de fortes chances qu'un événement particulier soit survenu : maladie, stress, chaleurs, mise-bas. Le changement s'opère avant que les soigneurs repèrent l'évènement. Ensuite, nous avons appliqué le random forest à des attributs décrivant les séries d'activité de 24 h (minimum, maximum, moyenne, autocorrélations…). Nous classons correctement 99% des séries témoins (= quasiment pas de fausse alertes) et dans un épisode qui entoure une maladie ou un événement reproductif, la probabilité de classer correctement au moins une série varie de 94 à 100%. L'apprentissage automatique appliqué à des séries temporelles semble donc un outil très puissant pour analyser le comportement des animaux et diagnostiquer son état interne.

Machine Learning applied to behaviour monitoring to detect diseases, reproductive events and disturbances in dairy cows

VEISSIER I. (1), LARDY R. (1), MIALON M.M. (1), RUIN Q. (1), ANTOINE V. (2), KOKO J. (2)
(1) Université Clermont Auvergne, INRAE, VetAgro Sup, UMR Herbivores, 63122 Saint-Genès-Champanelle, France

SUMMARY
Animal behaviour is very sensitive to any external disturbance or change in the animal internal state. We sought to predict and classify a wide range of health, stress and physiological states from daily activity of dairy cow. We had six datasets (total, 122,000 cows*days). The caretakers noted the condition of the cows: illness, oestrus, calving, disturbances (handling, mixing. Cows were equipped with sensors to estimate the duration of activities - eating, resting, in the alleys - from which the activity level was calculated. We modelled the rhythm of activity by the Fourier transform (Harmonics 0 and 1). If the difference between the models obtained during 2 series of 24 h is higher than a certain threshold, there is a strong chance that a particular event has occurred: illness, stress, heat, farrowing. The change takes place before the caretakers spot the event. Next, we applied random forest to attributes describing the 24 h activity series (minimum, maximum, average, autocorrelations...). We correctly classify 99% of the control series (= very rare false positive) and in an episode surrounding a disease or a reproductive event, the probability of correctly classifying at least one series varies from 94 to 100%. Machine learning applied to time series seems therefore a very powerful tool to analyse the behaviour of animals and diagnose its internal state.

INTRODUCTION

Le comportement d'un individu est très sensible à toute perturbation extérieure ou modification de l'état interne. Ainsi, des animaux stressés peuvent être plus actifs que la normale. Une vache cherche à s'isoler et est agitée à l'approche de la mise-bas (Jensen 2012). L'agitation accompagne également les chaleurs (Reith et Hoy 2018). Enfin, un animal malade diminue son activité, dort plus et à des moments où il est normalement éveillé, réduit la prise d'aliment et d'eau) et interagit moins avec son environnement (Dantzer et Kelley 2007 ; De Boyer Des Roches et al., 2017). Ces modifications peuvent apparaitre avant les signes cliniques. Les variations d'activité au cours de la journée, en particulier l'alternance jour-nuit - semblent des indicateurs particulièrement sensibles et précoces d'un trouble (Veissier et al., 1989). Ainsi des taurillons atteints d'une affection respiratoire sont moins actifs leur activité varie moins au cours de la journée, 3-6 jours avant l'apparition des signes cliniques (Marchesini et al., 2018).
Des changements subtils de comportement, comme la modification des rythmes (par exemple, le rythme journalier de veille/sommeil), ne peuvent pas être repérés par des observations ponctuelles. Leur détection ne peut se faire que si l'activité des animaux est enregistrée en continu, ce qui est maintenant possible grâce à des capteurs (accéléromètres, capteurs de positionnement, lecteurs RFID pour connaitre la

proximité des animaux avec une zone d'intérêt, …)(Buller et al., 2020).
Les systèmes développés pour l'élevage de précision utilisent de tels capteurs pour détecter quand les animaux sont dans des états spécifiques tels qu'en chaleurs, prêts à mettre bas, malades, … Le plus souvent ils ne détectent qu'un trouble à la fois ou plusieurs mais sans les distinguer (Stachowicz et Umstätter 2021). Nous avons cherché à prédire et à classer un large éventail d'états de santé, de stress et d'états physiologiques à partir de l'activité quotidienne des vaches, et plus particulièrement à partir d'indicateurs de rythme d'activité.

1. MATERIEL ET METHODES

Nous rapportons des travaux conduits sur plusieurs troupeaux de vaches Holstein, maintenus à l'intérieur, en fermes commerciales ou expérimentales (INRAE Herbipôle, IRTA). Nous avons utilisé 6 jeux de données comprenant

- des informations sur l'état des vaches (mammite, boiterie, acidose, vêlage, œstrus,…) ou des événements extérieurs potentiellement stressant (par ex changement de parc, manipulations) enregistrées dans les cahiers d'élevage par les soigneurs (et pour un jeu de données, des informations sur le pH ruminal afin d'identifier l'acidose)
- des données d'activité enregistrées à l'aide de capteurs.

Pour cinq jeux de données, l'activité était fournie par le système de localisation en temps réel CowView® (GEA Farm Technologies) : les vaches étaient équipées d'un tag émettant des ondes captées par des antennes placées au plafond, la position de l'animal est déterminée par triangulation, l'activité est inférée de la position (se repose si la vache est dans la logette, mange si elle est près de l'auge, sinon debout dans l'aire d'exercice). Pour un jeu de données, les vaches étaient équipées d'accéléromètres HOBO (Pendant G Acceleration Data Loggers, Onset Computer Corp) déterminant la position couché / debout et leurs auges étaient munies de portillons électroniques déterminant le temps passé à manger.
Nous calculons le niveau d'activité d'une vache pour chaque heure de la journée selon l'équation :
Niveau d'activité = -0.23 (temps passé au repos) + 0.16 (temps passé debout sans manger) + 0.42 (temps passé à manger)
Les poids ont été définis à partir d'une analyse factorielle des correspondances (Veissier et al 2001). Dans la suite de l'article, le terme « activité » réfère au niveau d'activité ainsi calculé.

2. MISE EN EVIDENCE DES MODIFICATIONS DE RYTHME D'ACTIVITE

Dans un premier temps (Veissier *et al.,* 2017), nous avons utilisé deux indicateurs de l'activité :

- l'activité moyenne du jour
- l'écart-type entre heures du jour, reflétant les variations intra-jour

calculés pour chaque jour et vache dans un jeu de données comprenant 350 vaches observées pendant 5 mois. Nous avons conduit une analyse de variance pour données répétées pour étudier la relation entre ces deux indicateurs et l'état des vaches : mammite, boiterie, autre maladie, en chaleurs. D'autres problèmes de santé ont été diagnostiqués mais avec une occurrence trop faible, ne permettant pas de les inclure dans l'analyse (10 cas de métrite, 5 d'acétonémie et 2 de colique). Les vaches étaient plus actives et les variations intra-jour étaient moins marquées le jour où une mammite ou des chaleurs étaient relevées. Les mêmes modifications étaient observées la veille de chaleurs mais pas l'avant-veille où les variations intra-jour étaient au contraire plus marquées que la normale. Les deux jours précédant l'observation d'une mammite, les vaches présentaient des variations intra-jour moins marquées. Lors de boiterie, le niveau moyen d'activité ne change pas mais les variations intra-jours sont moins marquées le jour même de l'observation et la veille (Tableau 1). Ces résultats ont confirmé que le rythme d'activité d'une vache est modifié lorsqu'elle est malade ou en chaleurs et que ces modifications précèdent les signes cliniques.

	Chaleurs	Mammite	Boiterie
Jour 0			
Niveau	+	+	=
Variations	-	-	-
Jour -1			
Niveau	+	=	=
Variations	-	-	-
Jour -2			
Niveau	+	=	=
Variations	+	-	=

Tableau 1 Mise en évidence de modifications d'activité selon l'état des vaches (niveau moyen et variations entre heures de la journée)

3. DETECTION DES CHANGEMENTS DE RYTHME D'ACTIVITE

Dans un second temps (Wagner *et al.,* 2021), nous avons cherché à modéliser la courbe journalière d'activité des vaches afin de détecter des changements. Nous avons utilisé 4 jeux de données différents du premier (total = 119000 jours x vaches). Les jeux ont été décomposés en séries de 36 h. Nous avons appliqué la transformation de Fourier sur les 24 premières et les 24 dernières heures de chaque série. Nous avons retenu la fréquence fondamentale et la première harmonique. Ainsi chaque série de 24 h est décrite par un modèle défini par l'activité moyenne et une sinusoïde représentant la cyclicité sur 24 h (en d'autres termes, le rythme circadien). Nous calculons ensuite la distance entre les deux modèles représentant les premières et dernières 24 h de la série de 36 h (Figure 1).
Un épisode de 4 jours a été défini autour d'une maladie, d'un vêlage ou de chaleurs : le jour de l'observation, deux jours avant, le jour suivant sont labélisés comme « vache dans un état particulier ». La semaine suivante est exclue de l'analyse. Le reste des jours sont considérés comme des jours normaux.
Pour chaque jeu de données, nous avons défini la distance au-delà de laquelle nous considérons que l'activité a changé au sein d'une série de 36 h. Ce seuil a été optimisé de sorte que les vaches qui ont été notées malades, en chaleurs ou vêlant par les soigneurs soient effectivement considérées comme ayant changé leur activité. Ces opérations ont été conduites sur un tiers de chaque jeu de données tiré au hasard. Nous avons ensuite testé ce seuil sur le reste des données. Avec cette méthode, près de 80% des jours notés sans problème sont détectés sans changement d'activité. Seulement 30% des jours notés avec une maladie, des chaleurs ou un événement potentiellement stressant, sont détectés avec changement d'activité. Or lors d'un épisode de maladie, vêlage ou chaleurs, la recherche de changement de rythme d'activité s'effectue dès deux jours avant que cet état soit observé par les soigneurs et se poursuit jusqu'au lendemain de l'observation. Si on s'intéresse aux épisodes où au moins une série de l'épisode est détectée avec un changement, la performance atteint 95% de détection. La méthode (appelée FBAT pour Fourier-Based Approximation with Thresholding) permet donc d'identifier les jours x vaches où un état particulier de la vache est suspecté : maladie, vêlage, œstrus, stress. Mais elle ne permet pas de discriminer ces états.

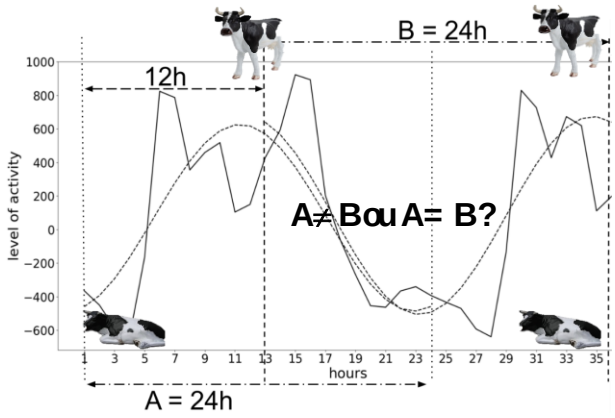


Figure 1 Modélisation de l'activité par la transformée de Fourier. Trait plein : niveau d'activité mesuré pendant 36 h ; traits pointillés : niveau d'activité modélisé pour deux séries A et B de 24 h décalées de 12 h.

4. IDENTIFICATION DES MALADIES, DES ETATS REPRODUCTIFS ET DU STRESS

Dans un troisième temps, nous avons décrit plus finement chaque série de 24 h à l'aide de 32 attributs statistiques afin de décrire

- l'activité globale (moyenne, médiane, minimum, maximum, quantiles 10-90 et 25-75
- les variations entre heures (écart-type, différences successives,…)
- la cyclicité des variations (autocorrélations d'ordre 1 à 11, harmoniques de Fourier 1 à 4).

Ces calculs ont été appliqués à 5 jeux de données (dont les 4 précédents, 121000 jours x vaches).

Nous avons appliqué un algorithme d'apprentissage automatique (Machine Learning) pour classer les séries de 24 h selon l'état des vaches. Nous avons appliqué l'algorithme Random Forest (1000 arbres construits, 10 tirages aléatoires des jeux d'apprentissage et de test pour chaque jeu de données) (Figure 2) Les attributs ayant un poids supérieur à 4% dans la classification étaient: la moyenne ; la moyenne quadratique, l'écart-type ;la moyenne quadratique des différences successives (RMSSD) ; le maximum ; les quantiles 10, les 25, 50, 75, 90 % ; les indices Kurtosis et Skewness ; harmoniques de Fourier 1 à 4 ; les autocorrélations d'ordre 1, 2, 3, 4, 5. Nous avons classé correctement 99% des séries témoins (c'est-à-dire sans état particulier d'une vache) et 42% des séries labélisées avec un état particulier. Dans un épisode qui entoure une maladie ou un événement reproductif, la probabilité de classer correctement au moins une série variait de 94 à 100% pour les vêlages, les œstrus, les boiteries, les mammites, l'acidose ruminale, les autres maladies ; elle était de 84% pour les mélanges d'animaux. Au cours d'un épisode, les séries non correctement classées sont généralement confondues avec des séries témoins et plus rarement avec des perturbations ou des acidoses, ce qui limite l'impact des confusions.

CONCLUSION

La répartition des activités au cours de la journée semble très sensible à l'état interne de l'animal. L'apprentissage automatique sur les séries temporelles semble un outil très puissant pour détecter et distinguer divers désordres de santé, les chaleurs, les vêlages, et le stress des vaches. Nous avons fait la preuve de concept qu'il est possible d'aller plus loin que les outils actuels d'élevage de précision en orientant vers un diagnostic afin de faciliter la prise de décision. Nous devons maintenant vérifier que les algorithmes utilisés sont généralisables à d'autres contextes.

Ces travaux ont bénéficié du soutien des projets européens #311825 EU-PLF et #730924 Smartcow et du programme 16-IDEX-0001 (CAP 20-25).

Reith S., Hoy S., 2018. Animal, 12, 398-407.
Stachowicz J., Umstätter C., 2021. Proc. R. Soc. B, 288, 20210190.
Veissier I., Le Neindre P., Trillat G., 1989. Applied Animal Behaviour Science, 22, 1-12.
Veissier I., Mialon M.-M., Sloth K.H., 2017. Journal of Dairy Science, 100, 3969-3974.
Wagner N., Mialon M., Sloth K., Lardy R., Ledoux D., M. S., De Boyer Des Roches A., Veissier I., 2021. Methods, 186, 14-21.

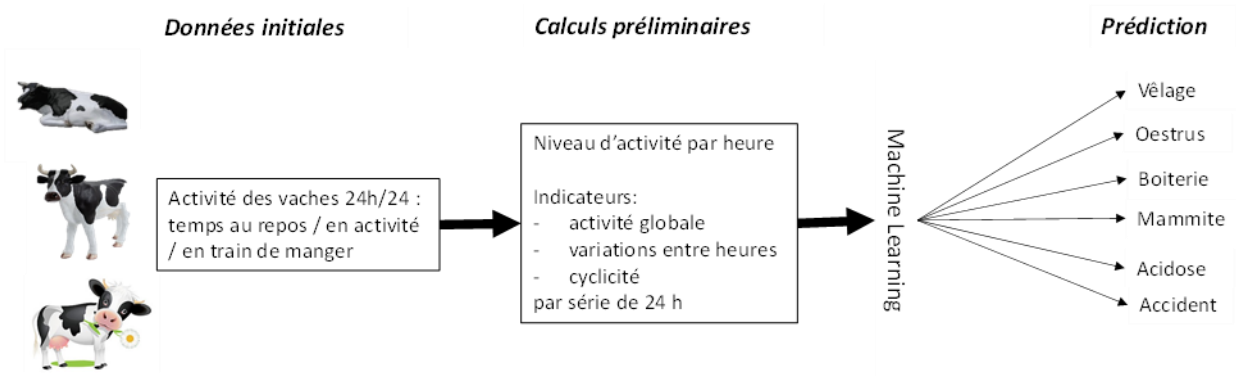


Figure 2 Développement du modèle de prédiction de l'état de la vache. Le modèle est calibré sur 10 à 30% du jeu de données (selon la taille du jeu) et testé sur le reste.