

Traçabilité des vaches laitières au pâturage grâce aux technologies de géolocalisation

ALLAIN C. (1), CHARPENTIER C. (1), PHILIBERT A. (2), LONIS W. (3), FISCHER A. (4).

(1) Institut de l'Élevage, Monvoisin, 35650 Le Rheu, France

(2) Institut de l'Élevage, 149 rue de Bercy, 75595 Paris, France

(3) Chambre d'agriculture de l'Oise, Rue Frère Gagne, 60021 Beauvais, France

(4) Institut de l'Élevage, 42 Rue Georges Morel, 49070 Beaucouzé, France

RESUME

Depuis 2007, plusieurs initiatives en Europe ont conduit à la démarche « lait de pâturage » pour répondre à la demande des consommateurs. En France, le référentiel impose que les vaches laitières (VL) passent au moins 6 h/j pendant au moins 120 j/an au pâturage. Cet essai visait à développer et tester un algorithme pour estimer le temps passé par les VL à l'extérieur du bâtiment d'élevage (TE) grâce à des capteurs GPS embarqués. Les données sont issues de 9 VL équipées d'un GPS parmi un troupeau de 70 vaches conduit en pâturage tournant et robot de traite pendant 56 jours. Le TE de référence a été estimé par une antenne RFID à l'entrée du bâtiment permettant de connaître les heures exactes de passages des animaux. L'algorithme développé permettait de détecter automatiquement le bâtiment d'élevage par une méthode de clustering, puis d'estimer le temps passé par les animaux à l'extérieur de ce bâtiment. L'estimation du TE quotidien individuel des vaches équipées était très corrélée à la référence ($R^2 = 0,79$) avec une erreur de prédiction de 91 min. Par rapport à la limite des 6h/j, la méthode développée a permis un taux de bonne prédiction de 93 %. Néanmoins, le TE moyen sur la période de l'essai a été surestimé en moyenne de 67 min notamment à cause d'une plus faible précision du GPS dans le bâtiment. Si l'effet du nombre de capteurs utilisés sur la capacité à détecter le bâtiment et sur l'estimation du TE moyen sur l'ensemble de la période n'était pas significatif, la grande variabilité du TE entre les VL a montré un risque plus élevé d'avoir une estimation biaisée pour un faible nombre de capteurs utilisés. L'estimation du TE grâce au GPS semble prometteuse pour objectiver la démarche « lait de pâturage ».

Traceability of dairy cows on pasture thanks to geotracking technologies

ALLAIN C. (1), CHARPENTIER C. (1), PHILIBERT A. (2), LONIS W. (3), FISCHER A. (4).

(1) Institut de l'Élevage, Monvoisin, 35650 Le Rheu, France

(2) Institut de l'Élevage, 149 rue de Bercy, 75595 Paris, France

(3) Chambre d'agriculture de l'Oise, Rue Frère Gagne, 60021 Beauvais, France

(4) Institut de l'Élevage, 42 Rue Georges Morel, 49070 Beaucouzé, France

SUMMARY

Since 2007, several initiatives in Europe have led to the "pasture milk" initiative to meet consumers demand. In France, the standards require that dairy cows (VL) spend at least 6 hours a day for at least 120 days a year on pasture. The aim of this trial was to develop and test an algorithm to estimate the time spent by the cows outside the barns (TE) using embedded GPS sensors. The data came from 9 cows equipped for 56 days with a GPS among a herd of 70 animals. The cows were milked with a robot and had access to pasture. The reference TE was estimated by a RFID antenna at the entrance of the building to determine the real exit timing of the cows. The developed algorithm automatically detected the barn by a clustering method and estimated the time spent by the animals outside of the barn. The estimation of the individual daily TE of equipped cows was highly correlated with the reference ($R^2 = 0.79$) with a prediction error of 91 min. Compared to the 6h/d limit, the method developed had a good prediction rate of 93%. Nevertheless, the average TE over the test period was overestimated on average by 67 min, due to the lower accuracy of the GPS inside the buildings. The effect of the number of sensors used on the ability to detect the building and on the estimation of the average TE over the whole period was not significant. However, the great variability of the TE between the cows showed a greater risk to have a biased estimation for a low number of sensors used. The estimation of the TE using GPS seems promising to objectify the "pasture milk" initiative.

INTRODUCTION

Depuis quelques années, l'exigence des consommateurs vis-à-vis des produits d'origine animale a fortement augmenté. Au-delà de la provenance et de la qualité de ces produits, ils souhaitent qu'ils soient respectueux de l'environnement et du bien-être animal (Marguet & Debéthune, 2018). Les consommateurs réclament ainsi de la transparence et de l'information sur ce qu'ils achètent (Frewer et al., 2005). En ce qui concerne les produits laitiers, le pâturage est incontournable aux yeux des consommateurs. Ils y voient non seulement un gage de bien-être, mais aussi d'autres atouts comme une alimentation naturelle ou une contribution à la biodiversité. Ce manque de transparence a contribué au déclin de la consommation de lait en France. Pour la relancer, les professionnels ont choisi de segmenter le marché (Marguet & Debéthune, 2018). Dès 2016, les

premiers « laits de pâturage » sont apparus dans les rayons français. Ce sont des laits dont les cahiers des charges contiennent un minimum obligatoire de jours pâturés. S'il existe différents cahiers des charges selon les laiteries, le CNIEL a récemment défini un référentiel précisant les règles minimales à respecter. Ainsi, pour qu'un lait puisse être qualifié de lait de pâturage, 95 % des vaches en lactation doivent avoir accès à au moins 10 ares/vache, à moins de 2 km du lieu de traite et au minimum pendant 120 j/an à raison de 6h/j (Rouyer, 2020). Actuellement, les contrôles reposent en grande partie sur des déclarations d'éleveurs (Marguet & Debéthune, 2018). Des audits d'élevage par des organismes certificateurs permettent aussi de prouver que les vaches pâturent, mais le contrôle du nombre de jours est plus arbitraire. Pour gagner la confiance du consommateur, les protocoles de certification doivent être fiables et réalisés en continu. Pour cela, le recours aux nouvelles technologies

peut permettre de réaliser des mesures objectives et en continu. Une collaboration initiée en 2019 avec la start-up Applifarm puis des travaux poursuivis avec l'entreprise Digitanimal dans le cadre du projet européen CattleChain 4.0, ont permis de développer un algorithme de traçabilité des vaches laitières au pâturage, grâce à des capteurs GPS embarqués sur les animaux.

1. MATERIEL ET METHODES

1.1. SCHEMA EXPERIMENTAL

L'essai a été mené à la ferme expérimentale de Derval (CA Pays De la Loire). L'exploitation possède 70 vaches laitières (VL) en traite robotisée et en pâturage tournant sur une parcelle divisée en 7 paddocks. Les vaches sont indépendantes et circulent soit librement entre le bâtiment et la pâture entre 8h00 et 17h00, soit sont restreintes au bâtiment entre 17h00 et 20h00 ou à la pâture sur le reste de la journée. Le temps passé par les VL à l'extérieur du bâtiment d'élevage (TE) n'est donc pas le même pour toutes les vaches chaque jour. Un unique couloir permet de relier le bâtiment et la pâture. Ce couloir est traversé dans sa largeur par une antenne RFID localisée dans le sol, détectant le passage des animaux par l'intermédiaire de leur podomètre. Associée aux données des portes de tri du bâtiment, ce système permet de connaître les heures réelles passées par les VL dans le bâtiment et donc au pâturage.

L'essai a duré deux mois, du 2 avril au 27 mai 2019. Parmi les 70 VL du troupeau, 9 ont été choisies de manière à représenter le troupeau en termes de rang et de stade de lactation. Elles ont toutes été équipées d'un collier GPS (Digitanimal, Espagne) et d'un podomètre (Nedap, Pays-Bas). Dans le cadre du projet, une fréquence de mesure des GPS de 11 min a été utilisée, équivalent à 130 données par jour dans le cadre d'un suivi quotidien continu. Les données de géolocalisation sont transmises sur un serveur via un réseau bas-débit économe en énergie proposé par Sigfox (Labège, France).

Deux méthodes ont ensuite été utilisées pour calculer le TE quotidien et individuel de chacune des vaches équipées. D'une part, ce temps a été calculé à partir des données de géolocalisation. Ce temps a ensuite été comparé au TE de référence fourni par les données de l'antenne.

1.2. ALGORITHME

L'objectif de l'algorithme est d'arriver à détecter le bâtiment d'élevage parmi les géolocalisations observées pour en déduire les géolocalisations en dehors du bâtiment, c'est-à-dire celles qui sont au pâturage et serviront pour estimer le TE. La détection du bâtiment d'élevage a été réalisée par une méthode de « clustering » avec le package Dbscan (Hahsler et al., 2019) du logiciel R (R Core Team, 2020). Cette méthode repose sur la détection d'un agglomérat de données, ici des géolocalisations, à partir de la densité des données. La densité des géolocalisations étant plus élevée dans le bâtiment par rapport aux pâtures, cette méthode devrait détecter sans problème un agglomérat correspondant au bâtiment d'élevage (Figure 1). Cet agglomérat a été détecté sur les données cumulées par période de 7 jours consécutifs glissant sur l'ensemble de la période. Chaque nouvel agglomérat est fusionné avec le précédent. Un polygone représentant l'agglomérat final est obtenu à l'issue de la période de l'essai. Les points en dehors du polygone sont utilisés pour estimer le TE.

Cet algorithme a été appliqué sur le jeu de données de géolocalisation des 9 vaches équipées d'un capteur GPS et d'un podomètre d'identification. Les données GPS manquantes liées à une absence de communication avec le réseau Sigfox ou aberrantes car largement en dehors de la zone pâturée ou du bâtiment ont été interpolées linéairement pour des trous de données allant jusqu'à 2 données consécutives manquantes. Au-delà, l'hypothèse de linéarité de

la géolocalisation est plus difficile à admettre. Seules les journées ayant au moins 110 données par jour après interpolation linéaire ont été conservées pour l'analyse de données. Par ailleurs, seul les journées communes aux données de géolocalisation et aux données de TE issues de l'antenne de référence ont été conservées pour la comparaison du TE estimé et de référence.

Ce nombre total de données par jour étant souvent inférieur à 130, le TE a ainsi été calculé comme suit pour le rapporter à une journée de 24h selon l'équation ci-dessous :

$$\text{TE moy. estimé (min/j)} = \frac{(\text{Nb de données/j hors du bât.} \times 11 \text{ min}) \times 1440 \text{ min}}{\text{Nb de données tot./j} \times 11 \text{ min}}$$

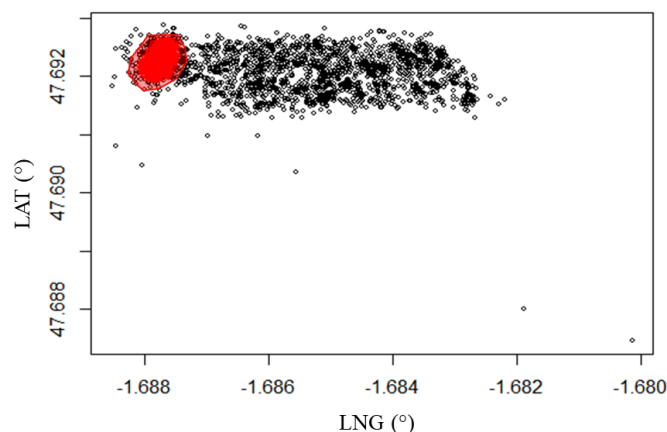


Figure 1 Détection du bâtiment (agglomérat et polygone) à partir de la méthode DBSCAN.

1.3. ANALYSE DES DONNEES

Compte tenu du coût des GPS à l'unité, quelques vaches uniquement dans chaque troupeau seront équipées. Ainsi, des tests ont été effectués afin d'estimer l'impact du nombre de colliers sur la capacité à détecter le bâtiment et sur le calcul du temps d'accès à l'extérieur. Pour cela, toutes les combinaisons de 2 à 9 capteurs ont été générées sous R (R Core Team, 2020) grâce à la fonction `combn` du package de base. Pour chaque combinaison de colliers, plusieurs paramètres ont été estimés : les coordonnées du centre de gravité et la surface de l'agglomérat du bâtiment détecté, le TE moyen estimé sur la période de l'essai ainsi que son écart-type inter-jours, et le TE moyen quotidien minimum obtenu sur la période. Une analyse de la variance avec les tests de Games-Howell a été réalisée pour tester l'effet du nombre de capteurs sur les différentes variables précédemment citées. Les tests post-hoc ont été réalisés avec la fonction `oneway` du package `userfriendlyscience` (Peters, 2017).

Le TE obtenu avec l'algorithme a été comparé au TE obtenu avec la méthode de référence. Seules les journées et les capteurs en communs aux deux jeux de données ont été conservés pour l'analyse. Une régression linéaire simple a été réalisée sous R avec la fonction `lm` du package de base de R (R Core Team, 2020). L'erreur de prédiction du modèle a été extraite et décomposée en biais moyen, biais de pente et en erreur aléatoire d'après la définition de Wallach et al. (2006). Pour détecter si l'estimation des TE réels et des TE estimés était en accord sur la limite des 6h/j, des calculs de spécificité et de sensibilité ont été réalisés. Pour cela, il a été considéré que le vrai positif était lorsque les TE réels et estimés étaient tous les deux inférieurs à 6h/j et le faux négatif lorsque le TE réel était de moins de 6h/j alors que le TE estimé était de plus de 6h/j.

2. RESULTATS

2.1. IMPACT DU NOMBRE DE COLLIERS SUR LA DETECTION DU BATIMENT D'ELEVAGE

Les caractéristiques moyennes du bâtiment d'élevage détecté par l'algorithme fluctuent très peu en fonction du nombre de colliers utilisés. En effet, il n'y a ni d'effets significatifs sur les coordonnées moyennes du centre de gravité, ni sur la surface moyenne du bâtiment détecté ni sur le TE moyen estimé sur la période ($p > 0,05$) (Figure 2). Toutes combinaisons confondues, la surface moyenne obtenue est de $3424 \pm 574 \text{ m}^2$ et est 4,1 fois plus grande que la surface réellement accessible par les vaches (Figure 2). Le TE moyen estimé durant l'essai pour l'ensemble des combinaisons était de 707 ± 2 min par jour.

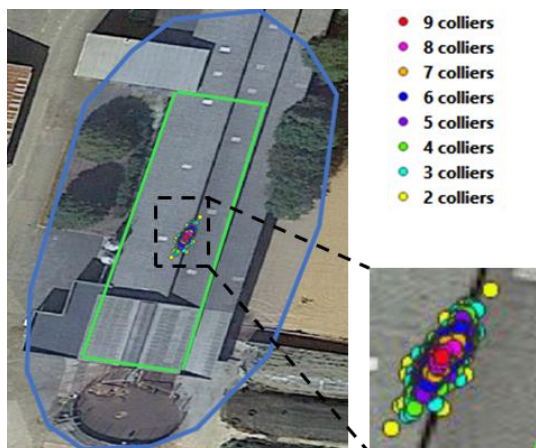


Figure 2 Centres de gravité pour les différentes combinaisons de colliers GPS (points de couleur) détectés par l'algorithme, surface réelle de la zone d'accès des vaches au bâtiment (rectangle vert) et polygone pour la combinaison de 9 colliers obtenu avec la méthode DBSCAN (bleu).

Néanmoins, plus le nombre de colliers utilisés était faible, plus les coordonnées du centre de gravité, la surface du bâtiment estimée et par conséquent le TE moyen estimé étaient variables. Ainsi, l'écart-type inter-combinaisons du TE passait de 173 ± 20 min pour les combinaisons de 2 colliers à 117 ± 3 min avec 8 colliers (Figure 3). De 2 à 6 capteurs, cette variabilité inter-combinaisons était significativement différente de la variabilité observée avec 7 ou 8 capteurs ($p < 0,05$). Par ailleurs, plus le nombre de colliers utilisé était élevé, plus le TE minimum obtenu sur la période de l'essai augmentait (Figure 4). A partir de 6 colliers les temps n'étaient plus significativement différents du TE minimum obtenu avec 8 colliers. 2/3 des combinaisons de 2 colliers et la moitié des combinaisons de 3 colliers avaient un TE minimum inférieur à 6h/j.

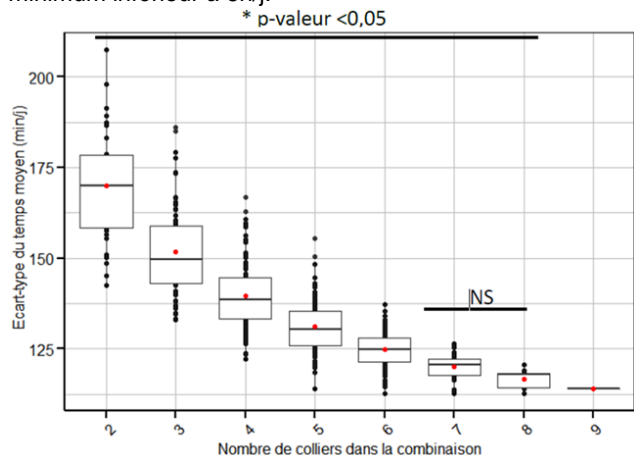


Figure 3 Effet du nombre de colliers sur l'écart-type du TE moyen. NS = non significatif.

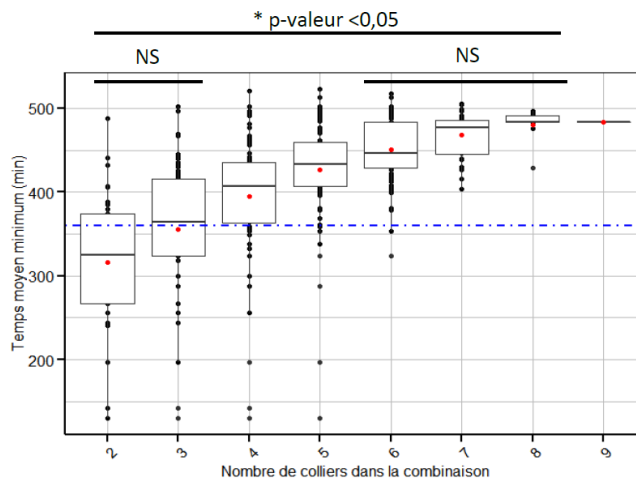


Figure 4 Evolution du TE moyen minimum sur les 56 jours d'essai en fonction du nombre de colliers utilisés dans la combinaison. La ligne bleue représente la limite des 6h/j. NS = non significatif.

2.2. ESTIMATION DU TEMPS EN EXTERIEUR

En utilisant les 9 colliers disponibles, le TE moyen estimé était de 734 ± 200 min et le TE moyen de référence de 667 ± 220 min. La régression linéaire des TE de référence et des TE estimés par vache et par jour indiquait une corrélation positive avec un coefficient de détermination de 0,79 (Figure 5). L'erreur moyenne de la méthode GPS était de 91 min dont 9% était expliqué par un biais de pente et 28% par un biais moyen. Cela indique que le TE estimé est en moyenne surestimé, mais que cette surestimation est plus importante pour des temps de pâturage faibles. Lorsque l'on compare les TE moyens par jour de l'ensemble des colliers utilisés, le coefficient de corrélation de détermination était de 0,78 et l'erreur moyenne de 67 min, dont 7% était expliqué par un biais de pente et 37% par un biais moyen.

La méthode GPS a surestimé le TE de référence dans 80% des cas, dont 50% à moins de 1h, 25% entre 1h et 2h, 14% entre 2h et 3h et 11% à plus de 3h. La validité intrinsèque de la méthode en utilisant les données de toutes les vaches par jour montre une sensibilité de 35%, une spécificité de 99% et un taux de bonnes prédictions de 93%. Lorsque les données moyennées par jour sont analysées, ces mêmes paramètres étaient respectivement à 50%, 100% et 96%.

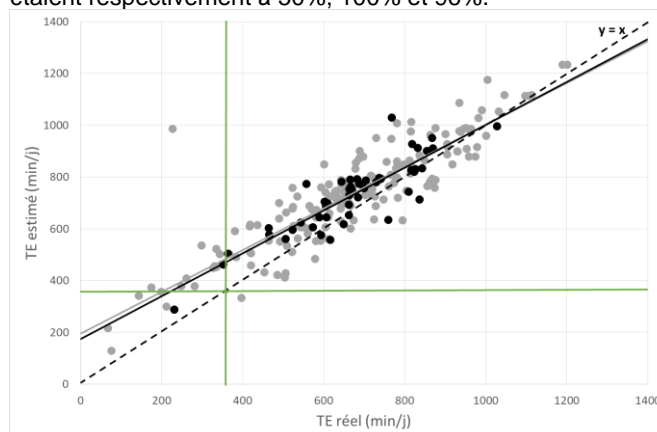


Figure 5 Régression linéaire du TE estimé par jour et par vache ($N = 176$; points gris) et du TE moyen par jour ($N = 50$; points noirs) en fonction du TE réel. La ligne pleine grise représente la droite de régression pour les points gris ($y=0,81x+195$; $R^2=0,79$) et la ligne pleine noire représente la droite de régression pour les points noirs ($y=0,83x+173$; $R^2=0,78$). La limite des 6h/j est représentée par les lignes vertes.

3. DISCUSSION

3.1. DETECTION AUTOMATIQUE DU BATIMENT ET EFFET SUR LE TEMPS EN EXTERIEUR

Les résultats présentés précédemment montrent que l'algorithme fonctionne bien pour détecter le bâtiment et notamment son centre de gravité. En revanche, la surface estimée est en moyenne 4,1 fois supérieure à la surface réellement accessible aux vaches dans le bâtiment. Si les GPS ont une précision d'environ 10 m à l'extérieur, celle-ci tombe à 50 m lorsqu'ils sont en intérieur. En effet, l'altération du signal causé par les bâtiments génère une difficulté de transmission entre les GPS et les satellites (Bhattacharya et al., 2015). Ce phénomène a été mis en évidence dans notre essai notamment sur le créneau horaire entre 17 et 18 h où les vaches étaient exclusivement au bâtiment (Figure 6). En effet, de très nombreux points sont localisés en dehors du bâtiment réel, et 4% des points sont également en dehors du polygone détecté par l'algorithme. Ces points localisés à l'extérieur augmentent artificiellement le TE car ils sont considérés comme extérieurs au bâtiment, donc comme géolocalisation au pâturage. Compte-tenu de la plus forte imprécision des GPS au bâtiment, le risque de surestimer le TE est d'autant plus grand que les animaux passent de temps en bâtiment. La prise en compte de l'analyse des trajectoires passées et postérieures à chaque donnée de géolocalisation (Bhattacharya et al., 2015) permettrait de réduire le risque d'identifier à tort des géolocalisation à l'extérieur du bâtiment.

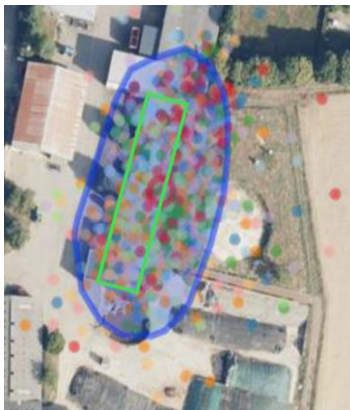


Figure 6 Points GPS générés entre 17 et 18h (heure à laquelle toutes les vaches sont en bâtiment) sur toute la période de l'essai. Le rectangle vert représente la surface d'accès réelle des vaches au bâtiment. Le polygone bleu représente le bâtiment détecté par DBSCAN.

3.2. CHOIX DU NOMBRE VACHES A EQUIPER

Les cahiers des charges des laits de pâturage étant variables selon les laiteries (6 à 8 h/j d'accès au pâturage pendant 90 à 200 j/an, Pruilh, 2019) et compte tenu du coût important des capteurs, la question du nombre d'animaux à équiper pour être représentatif du troupeau est primordiale. Par exemple, dans le cadre de notre essai, il existe une très grande variabilité des TE quotidien par vache en raison de l'accès libre et volontaire des vaches aux parcelles extérieures. Si le cahier des charges impose que sur la durée de pâturage (90-120 j/an) le troupeau doit passer en moyenne au moins 6 h/j à l'extérieur, alors 2 colliers suffiraient. En effet, dans notre cas, quelle que soit la combinaison de 2 colliers utilisée, le TE moyen sur la période de l'essai était toujours supérieure à 6 h/j. En revanche, si les exigences imposaient que sur cette période de pâturage, chacun des jours devait avoir un TE moyen supérieur à 6 h/j, il faudrait au minimum équiper 6 animaux. En effet, compte tenu de la grande variabilité que nous avons observée, 2/3 des combinaisons de 2 colliers testées et la moitié des combinaisons de 3 colliers testées ont eu au moins une journée avec un TE moyen quotidien inférieur à 6 h.

Une des pistes possible pour équiper plus d'animaux ou diminuer le coût de l'investissement consisterait à n'équiper que quelques animaux avec un capteur GPS et tous les autres animaux avec des boucles Bluetooth (Maroto-Molina et al., 2019). Les GPS communiquent avec les boucles Bluetooth en mesurant la distance qui les séparent et cela permet ainsi d'estimer la localisation de chaque animal pour un coût 10 fois moins élevé que des colliers GPS uniquement.

3.3. AUTRES PERSPECTIVES D'UTILISATION DES CAPTEURS EMBARQUES

Pour être massivement adoptés par les éleveurs laitiers, ces technologies doivent soit permettre d'améliorer significativement le prix du lait commercialisé, soit fournir d'autres services aux éleveurs. Parmi les services envisagés, l'automatisation du calendrier de pâturage utilisé par les éleveurs pourrait leur permettre de s'affranchir l'astreinte de son enregistrement. Actuellement, les calendriers de pâturage nécessitent la saisie de l'effectif exact qui pâture et de la chronologie de rotation des paddocks. En effet, grâce aux technologies de géolocalisation (GPS et boucles Bluetooth) embarquées sur les animaux, un suivi instantané des paddocks pâturés, des temps d'accès par paddock et de l'effectif exact pâturant pourra être mis en place (figure 7).

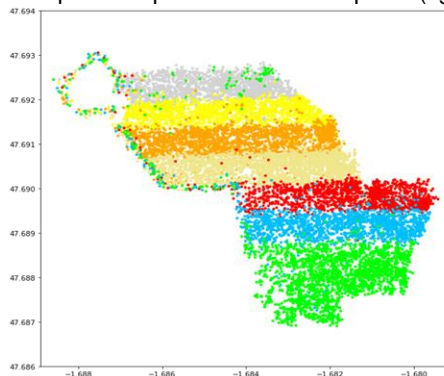


Figure 7 Visualisation des paddocks pâturés par les animaux sur la période de l'essai à partir des données GPS. Chaque couleur représente une journée différente.

CONCLUSION

La méthode utilisée semble prometteuse pour objectiver le temps d'accès à l'extérieur malgré la mise en évidence d'une surestimation et le besoin de la valider dans d'autres systèmes d'élevage. Toutefois, pour qu'il soit utilisé en routine et qu'il soit accepté par les éleveurs il sera nécessaire que cet outil puisse apporter d'autres services comme par exemple le remplissage automatique du calendrier de pâturage.

Les auteurs remercient le personnel de la ferme expérimentale de Derval pour la mise en place de l'essai ainsi qu'Applifarm, à l'initiative de ces travaux.

- Bhattacharya T., Kulik L., Bailey J., 2015. Pervasive and Mobile Computing, 19, 86-107.
- Frewer L. J., Kole A., Van de Kroon S. M. A., Lauwere C., 2005. Journal of Agricultural and Environmental Ethics, 18, 345-367.
- Hashler M., Piekenbrock M., Doran D., 2019. Journal of Statistical Software, 91, 1-30.
- Marguet M., Debéthune N., 2018. Terragricoles de Bretagne, 228.
- Maroto-Molina F., Navarro-Garica J., Principe-Aguirre K., Gomez-Maqueda I., Guerrero-Ginel J. E., Garrido-Varo A., Perez-Marin D. C., 2019. Sensors., 19, 2298.
- Peters G., 2017. R package version 0.7.2.
- Pruilh C., 2019. Réussir Lait, 331.
- R Core Team, 2020
- Rouyer B., 2020. Document interne, 6p.
- Wallach D., 2006. Elsevier, Ed. Amsterdam, 22-24