

Thèse de Doctorat

Pierre CLEMENT

Mémoire présenté en vue de l'obtention du

grade de Docteur d'Oniris - l'École Nationale Vétérinaire Agroalimentaire et de l'Alimentation Nantes-Atlantique

sous le label de L'Université Nantes Angers Le Mans

École doctorale : *Biologie-Santé Nantes-Angers*

Discipline : *Biologie, Médecine, Santé*

Spécialité : *Recherche clinique, innovation technologique, santé publique*

Unité de recherche : *UMR1300 Biologie, Epidémiologie et Analyse de Risque en santé animale (BioEpAR)*

Soutenue le 30 octobre 2014

Valeur informative de l'activité générale et de la durée de rumination pour la détection des troubles de santé et le phénotypage chez la vache laitière

JURY

Rapporteurs : **Didier CALAVAS**, Ingénieur de Recherche, ANSES de Lyon, France.
Philippe FAVERDIN, Directeur de Recherche, INRA de Rennes, France.

Examineurs : **Daniel BERCKMANS**, Professeur, Université Catholique de Louvain, Belgique.
Sébastien ASSIE, Maître de conférences, Oniris, Nantes, France.

Directeur de Thèse : **Nathalie BAREILLE**, Professeur, Oniris, Nantes, France

Co-encadrant de Thèse **Raphael GUATTEO**, Maître de conférences, Oniris, Nantes, France.

Avertissement de confidentialité

Ce document est confidentiel.

Sa diffusion est interdite jusqu'au **30 octobre 2016**.

Préambule

Le travail présenté dans cette thèse est issu de la collaboration entre l'entreprise Evolution (Rennes, France) et l'UMR Oniris-Inra 1300 « Biologie, Epidémiologie et Analyse de Risque en santé animale» (BioEpAR, Nantes, France). Il a été financé par l'entreprise Evolution (Rennes, France) à l'aide d'un Crédit Impôt Recherche dans le cadre de l'Institut Carnot en Santé Animale.

Les travaux de recherche ont été réalisés au sein de l'UMR1300 BioEpAR sous la direction de Mme Nathalie BAREILLE, Professeur à Oniris, entre octobre 2011 et octobre 2014. Le co-encadrement a été réalisé par M. Raphaël GUATTEO, maître de conférences à Oniris.

Les travaux de cette thèse ont également été guidés et réalisés grâce à différents partenaires dont l'Institut de l'Elevage et les chambres d'Agriculture du Maine et Loire et de Bretagne, via la participation des fermes expérimentales de Trevarez (29) et des Trinottières (49). Il a également été rendu possible grâce à l'implication d'éleveurs de vaches laitières ayant accepté de participer aux travaux.

**Merci aux différents partenaires qui ont permis la réalisation des travaux
présentés dans cette thèse.**

Remerciements

Tout d'abord un grand merci à Nathalie Bareille et Raphaël Guatteo pour m'avoir guidé, accompagné et corrigé au cours de ces 3 ans. Votre réactivité légendaire m'a souvent fait espérer (ou craindre !) que vous ayez corrigé mon travail dès le lendemain de ma sollicitation. Merci pour l'ambiance toujours constructive de nos brainstormings, et vos encouragements bienvenus lorsque je doutais. En espérant avoir acquis un peu de vos qualités à votre contact et avoir progressé au cours de ces 3 années.

Merci à Jean-Michel Philipot, Gilles Delaporte et Thomas de Bretagne pour avoir initié, accompagné ou suivi ce travail. C'était une belle occasion pour moi de découvrir la recherche en entreprise.

Merci aux chambres d'agriculture de Bretagne et du Maine et Loire et à l'Institut de l'Élevage pour avoir participé à ces travaux. En particulier, merci à Jean-Michel Lamy, Guylaine Trou, Julien François, Marylise Le Guenic, Audrey Chanvallon, Clément Allain et Benoit Rouillé pour avoir suivi et orienté ce travail, ainsi que pour la collecte des données. Un grand merci également au personnel des fermes expérimentales de Trevarez et des Trinottières pour le travail important réalisé pour la collecte des données nécessaires à cette thèse.

Merci aux éleveurs ayant participé aux études menées dans cette thèse, en particulier à Jean-Yves et Jacques Brard ainsi qu'à Christophe Dugast pour leur accueil chaleureux. En espérant que les résultats puissent leur être utiles à l'avenir.

Merci à René Quiniou pour l'apprentissage du CUSUM, sa disponibilité et sa gentillesse tout au long de ce travail, et pour m'avoir empêché de me perdre dans le labyrinthe de l'INRIA. Bonne continuation à toi. Un grand merci également à Anne Lehebel et Aurélien Madouasse pour m'avoir aiguillé dans le monde ardu des statistiques et des séries temporelles.

Je remercie les membres de mon comité de thèse, René Quiniou, Luc Delaby, Christian Ducrot, Jean-Michel Philipot et Christine Fourichon pour leurs conseils avisés pour faire aboutir ce travail.

Merci à Jean-Yves Audiart, Erwann Helleu, Michel Leroux et Didier Billon pour avoir su résoudre dans la bonne humeur tous les petits tracas du quotidien - informatiques, logistiques, bibliographiques ou concernant les bases de données - rencontrés. Merci pour votre disponibilité et votre gentillesse.

Merci à Evelyne Fâche, Sylvie Alonso et Juliette Bordot pour avoir toujours su trouver une solution aux petits écueils administratifs, alors même qu'il m'arrivait parfois d'être dans le gruppetto, voire franchement hors délai.

Merci à tous les autres membres de l'UMR Bioepar pour les discussions enrichissantes et votre disponibilité en cas de besoin.

Merci pour leur contribution aux stagiaires Médoune Fall et Thibault Desquesnes qui ont participé à ce travail.

Un énorme merci aux amis de Bioepar pour les moments partagés et pour votre soutien au cours de ces 3 années. Merci à Aurélien, Anne-Frieda, Julie, Mily, Manon, Simon, Anne, Arnaud, Amélie, Sandie, Rozenn, Eric, Elsa, Mohamed et Philippe pour les discussions enflammées, les cafés (ou les thés !) serrés, et les amitiés. Un merci spécial à la team « historique » du G4 : Nadine dont la bonne humeur égayait le bureau et Maïd pour sa simplicité et sa gentillesse. Quelle chance d'avoir partagé cette aventure humaine avec vous ! Et merci à tous les autres doctorants et post-doctorants de Bioepar pour faire vivre la solidarité et la bienveillance au quotidien.

Merci à ma famille, mes parents et beaux-parents, Yann et Nanou et Iwan et Angélique, Solène et Corentin pour m'avoir soutenu tout au long de l'aventure, et notamment dans les moments difficiles. Pour le temps que je n'ai parfois pas pu passer avec vous.

Enfin merci à Sabine pour les sacrifices consentis pour que je puisse réaliser cette thèse. Pour ton aide concrète et ton soutien de tous les instants. Je serai là pour toi pour la suite de notre vie.

Sommaire

Avertissement de confidentialité.....	i
Préambule	ii
Remerciements	iii
Liste des tableaux.....	xi
Liste des figures.....	xiii

Chapitre 1. Introduction générale 1

1. Contexte et enjeux du monitoring en élevage	2
2. Positionnement et problématique générale du travail de recherche	8
2.1 Objectifs et contexte du projet de thèse	8
2.2 Description du système de monitoring support de la thèse.....	9
2.3 Compétences scientifiques mobilisables autour du monitoring.....	10
2.4 Positionnement du travail de recherche.....	13
2.5 Problématique générale	14
3. Etat des connaissances et questions de recherche sur la valeur informative de l'activité générale et des durées de rumination	15
3.1 Qualité de la mesure de l'activité par Heatime-HR®	15
3.2 Qualité de la mesure des durées de rumination par Heatime-HR®	16
3.3 Intérêt de l'activité générale et des durées de rumination en termes de phénotypage	19
3.4 Valeur informative et intérêt pour la détection des chaleurs.....	21
3.5 Valeur informative et intérêt pour la détection des troubles de santé	24
3.5.1 Signification clinique des baisses d'activité générale et durée de rumination .	24
3.5.2 Peu d'études basées sur les mesures du Heatime-HR®	25

3.5.3 Points de vigilance méthodologiques pour évaluer la valeur informative d’alertes générées à partir de données issues de capteurs	27
3.5.3.1 Temporalité des évènements et attribution des alertes aux troubles de santé.....	27
3.5.3.2 Statistiques opérationnelles pour l’évaluation des alertes.....	30
3.5.3.3 Choix d’une méthode de référence pertinente pour la détection des troubles de santé	30
4. Objectif général, stratégie et questions scientifiques traitées, plan de thèse	33
4.1 Intérêt de la mesure de l’activité générale et des durées de rumination pour le phénotypage	34
4.2 Valeur informative d’alertes générées à partir de l’activité générale et des durées de rumination pour la détection des troubles de santé	34
Références Bibliographiques.....	36
Annexe du chapitre 1.....	43
Chapitre 2. Evaluation de l’intérêt de la mesure des durées de rumination pour la prédiction des quantités ingérées chez des vaches laitières en lactation	45
1. Résumé du chapitre 2	46
2. Interpretative summary	47
3. Abstract	48
4. Short communication: added value of rumination time for the prediction of dry matter intake in lactating dairy cows.....	48
5. Acknowledgments.....	55
References	56
Chapitre 3. Valeur informative des durées de rumination et de l’activité générale pour la détection des troubles de santé.....	59
Partie I. Intérêt du monitoring des baisses de durées de rumination ou d’activité générale pour la détection des troubles de santé chez les vaches laitières	61

1. Résumé de la Partie I.....	61
2. Interpretative summary	62
3. Abstract	63
4. Introduction	64
5. Materials and methods	65
6. Results	75
7. Discussion	83
8. Conclusions	87
9. Acknowledgments	88
References	89
Partie II. Intérêt du monitoring des baisses de durées de rumination et d'activité générale pour la détection des chutes de production laitière chez les vaches laitières	92
1. Résumé de la Partie II	92
2. Abstract	93
3. Introduction	94
4. Materials and methods	96
5. Results	103
6. Discussion	106
7. Conclusions	109
References	110
Partie III. Evaluation d'alertes de santé générées chez des vaches laitières à partir de données d'activité et de rumination issues de capteurs, par rapport à un examen clinique vétérinaire systématique.	114
1. Résumé.....	114
2. Introduction	115
3. Matériel et méthodes	116
3.1 Principe et design de l'étude	116
3.2 Echantillon d'étude	117

3.2.1 Critères de choix du troupeau	117
3.2.2 Animaux et conditions d'élevage	118
3.3 Algorithmes initiaux et production des alertes de santé.....	119
3.4 Examen clinique vétérinaire	121
3.4.1 Critères d'inclusion et d'exclusion et choix des vaches à examiner	121
3.4.2 Réalisation des examens cliniques vétérinaires.....	122
3.5 Stratégie d'analyse	124
3.5.1 Exposition aux alertes de santé optimisées.....	124
3.5.2 Analyse du statut de santé	125
3.5.3 Calcul des valeurs informatives des alertes de santé optimisées.....	125
4. Résultats	126
4.1 Statistiques descriptives	126
4.2 Valeur informative des alertes de santé optimisées	128
5. Discussion	129
6. Conclusion.....	134
Références bibliographiques	135
Annexes du chapitre 3-III.....	137
Chapitre 4. Discussion générale	142
1. Principaux résultats	144
1.1 La connaissance des durées de rumination ne permet pas de mieux prédire les quantités ingérées	144
1.2 Les alertes de santé développées ne se substituent pas à l'observation des troubles de santé par l'éleveur.....	145
1.3 Des alertes de santé plutôt précoces	150
1.4 Comment évaluer les outils de monitoring : quels évènements veut-on détecter ?	152
1.4.1 Les questionnements initiaux autour de la méthode de référence	152
1.4.2 Intérêts et limites des 3 méthodes de référence utilisées	154

1.4.3 Que conclure de la confrontation des résultats obtenus avec les 3 méthodes de référence utilisées ?	156
1.5 Complémentarité de l'activité et de la rumination pour la détection des troubles de santé.....	158
2. Généralisation des résultats	159
2.1 Une estimation de l'ordre de grandeur des valeurs informatives des alertes de santé	159
2.2 Généralisation à d'autres systèmes de production	162
2.3 Les algorithmes commerciaux n'ont pas été évalués	163
3. Originalités et limites de la stratégie d'analyse.....	164
3.1 Une évaluation d'emblée multi-maladies.....	164
3.2 La spécificité des alertes a été privilégiée dans un contexte de faible prévalence des troubles de santé	165
3.3 Les évènements physiologiques pouvant impacter activité et rumination ont été pris en compte dans l'analyse des données	167
3.4 Le CUSUM : une méthode robuste mais où la prédiction des valeurs attendues d'activité et de rumination s'est révélée difficile	167
4. Perspectives	168
4.1 Perspectives pour les éleveurs	168
4.2 Perspectives pour les organismes de développement agricole et les conseillers d'élevage	171
4.3 Perspectives pour les développeurs/évaluateurs d'outils de monitoring.....	174
4.3.1 Des outils de détection des troubles de santé perfectibles.....	174
4.3.2 Une méthodologie réutilisable pour l'évaluation d'autres outils.....	176
4.4 Perspectives pour la recherche	177
4.4.1 Proposer une méthode d'évaluation standardisée des outils de monitoring de la santé	177
4.4.2 Evaluer la valeur informative des alertes de santé à partir des résultats de plusieurs tests imparfaits	177

4.4.3 L'utilisation d'autres méthodes mathématiques et statistiques pourraient permettre de mieux détecter les troubles de santé	178
4.4.4 Des analyses complémentaires pourraient permettre de mieux évaluer les performances de Heatime-HR®	179
4.4.5 Utiliser les données issues de capteurs pour le phénotypage de caractères de santé.....	180
4.4.6 Utiliser les données issues de capteurs pour la surveillance syndromique.....	182
4.4.7 Etudier les conditions techniques et économiques d'un déploiement réussi des technologies de l'élevage de précision	182
Références bibliographiques	185
Conclusion générale	192
Liste des publications et communications réalisées et prévues.....	196
Annexe générale.....	198

Liste des tableaux

Chapitre 1

Tableau 1-1 : Panorama des principaux outils de monitoring disponibles et leurs applications.....	5
Tableau 1-2 : Synthèse des études de validation des durées de rumination mesurées par le système Heatime-HR®.....	16
Tableau 1-3 : Synthèse des valeurs informatives obtenues pour la détection des chaleurs à l'aide des systèmes Heatime ou Heatime-HR®.....	21
Tableau 1-4 : Statistiques relatives à la valeur informative d'un système d'alerte.....	31

Chapitre 2

Table 2-1 . Descriptive statistics (mean, standard deviation, minimum, maximum) of the variables included into the models (n = 371 cow weeks).....	52
Table 2-2 . Statistical values obtained for NRC and NRC _{RT} models fitted to weekly DMI of individual cows, and accuracy and precision of the 2 models.....	53

Chapitre 3

Partie I

Table 3-I-1 : Standard form used by farm staff for recording health disorders. Definition of the type and severity of the health disorders.....	67
Table 3-I-2 . Number of health disorders detected by farm staff over the study (Ig=Ignored for the analysis, E=Excluded from the analysis).....	76
Table 3-I-3 : Informative values of CUSUM detections applied on Rumination Time (RT), Neck Activity (NA) or (RT or NA) for 3 CUSUM settings, for the main health disorders detected by farm staff.....	80

Table 3-I-4. Predictive values of CUSUM on RT or NA detections for different HD prevalences.....	81
---	----

Partie II

Table 3-II-1: Standard form used by farm staff for recording health disorders. Definition of the type and severity of the health disorders.....	97
--	----

Table 3-II-2: Descriptive statistics of major drops in milk production detected in the study (n = 347).....	103
--	-----

Table 3-II-3: Number and severity of health disorders detected by farm staff concomitantly to a drop in milk production (between 5 d before and 5 d after) or independently (others).....	104
--	-----

Table 3-II-4: Results of the CUSUM on RT or NA (number of events) for the detection of the major drops in milk production (n = 288) and of the HD detected by farm staff (n = 416).....	105
--	-----

Partie III

Tableau 3-III-1 : Principe du calcul de la sensibilité (Se) et de la spécificité (Sp) des alertes de santé (a, b, c, d = effectifs).....	117
---	-----

Tableau 3-III-2 : Paramètres utilisés pour les algorithmes CUSUM initiaux et optimisés (k=constante du CUSUM, h=limite de décision du CUSUM, Cf Chapitre 3.I).....	120
---	-----

Tableau 3-III-3 : Principaux paramètres observés lors de l'ECV.....	123
--	-----

Tableau 3-III-4 : Description des maladies et anomalies diagnostiquées lors des examens cliniques vétérinaires (ECV).....	127
--	-----

Tableau 3-III-5 : Sensibilité des alertes de santé et intervalle de confiance à 95% (IC _{95%}) en fonction de la sévérité du trouble clinique.....	128
---	-----

Tableau 3-III-6 : Grille de notation des métrites en fonction des observations au vaginoscope.....	140
---	-----

Tableau 3-III-7 : Grille de notation des mammites.....	140
---	-----

Tableau 3-III-8 : Grille de notation des boiteries.....	141
--	-----

Liste des figures

Chapitre 1

Figure 1-1 : Principales mesures pouvant être effectuées sur les vaches laitières, localisation des capteurs associés et domaines d'utilisation (d'après Idele, 2012).....4

Figure 1-2 : Principe de la comptabilisation des évènements pour le calcul de la valeur informative des alertes générées par un système de monitoring (FP=faux positif, VN=vrai négatif, VP=vrai positif, FN=faux négatif).....29

Chapitre 2

Figure 2-1. Scatter plot representing average time spent ruminating (h/d) against average DMI (kg/d) over the trial (n = 41 cows).....54

Figure 2-2. Scatter plot representing residual DMI (kg/d) against centered predicted DMI (kg/d) obtained with NRC_{RT} model (n = 371 cow weeks).....55

Chapitre 3

Partie I

Figure 3-I-1. Example of rumination time measured for one cow around an health disorder (solid line). The reference rumination time (dotted line), computed as the mean rumination time measured for that cow on the same 2 h time block of the day during the past 7 d is also shown. One significant difference (horizontal arrow) between measured and reference rumination time was identified just after a mastitis detection (vertical arrow), using a method based on the cumulative sum test.....70

Figure 3-I-2. Principle of computation of events for sensitivity and specificity calculation (TN = True Negative. TP = True Positive. FP = False Positive. FN = False Negative).....73

Figure 3-I-3. Distribution of the time of first detection by the CUSUM on RT or on NA compared to the detection by farm staff for the main health disorders.....82

Partie II

Figure 3-II-1: Example of one loss characteristic calculation.....99

Figure 3-II-2: Time-lag between the detection by the CUSUM on RT or NA (n = 103) and the first day of the drop in milk production.....106

Chapitre 1. Introduction générale

1. Contexte et enjeux du monitoring en élevage

En 2009, la France comptait 79 000 exploitations bovines laitières mais le nombre d'exploitations a tendance à diminuer depuis au moins les années 1980. Parallèlement, le nombre de vaches par troupeau continue à augmenter rapidement sans augmentation de la main d'œuvre associée, en France comme ailleurs en Europe. Entre 1990 et 2010, la taille moyenne des troupeaux laitiers en France a doublé (45 vs 23 vaches) (Agreste, 2012). Le nombre de bovins par éleveur a augmenté fortement, notamment depuis les années 1990 (Agreste, 2000). La fin des quotas laitiers prévue en 2015 pourrait encore accentuer cette tendance à l'agrandissement des exploitations.

Afin de seconder ou remplacer l'éleveur pour certaines tâches quotidiennes, des systèmes automatisés ont donc été développés depuis une vingtaine d'années. Historiquement, ce sont les systèmes de traite automatisée, vendus depuis les années 1990, qui ont fait arriver les nouvelles technologies dans les élevages équipés. En 2009, plus de 8 000 exploitations utilisaient des robots de traite, dont plus de 90% étaient situées en Europe de l'ouest (de Koning, 2010). Ces systèmes ont nécessité le développement de capteurs et d'outils électroniques et informatiques pour assurer leur fonctionnement en autonomie, et notamment le contrôle automatisé de la qualité du lait issu de la traite, indispensable à sa commercialisation. Le lait de vaches ayant une mammite par exemple, impropre à la consommation, doit règlementairement être écarté de la consommation humaine (de Koning, 2010). Ainsi, la recherche a initialement porté sur le développement de systèmes aptes à détecter la majorité des mammites à partir de mesures automatisées sur le lait (conductivité électrique, couleur, température du lait, concentration en cellules somatiques...). L'automatisation de certaines tâches s'inscrit dans un mouvement plus large d'adoption des nouvelles technologies en agriculture depuis les années 1980. En effet, les nouvelles technologies apparaissent pour les agriculteurs comme un gain de temps et une réponse rationnelle à des obligations réglementaires (déclaration des naissances, traçabilité, déclarations relatives à la politique agricole commune par exemple). D'autres forces comme le régime fiscal agricole favorable à l'investissement technique, la puissance des techno-imaginaires qui font apparaître les nouvelles technologies comme des moyens de maîtriser l'espace (ubiquité) et le temps, ou encore l'importance des structures professionnelles agricoles et de conseil en élevage promotrices d'innovations, ont joué des rôles majeurs dans l'adoption des nouvelles technologies dans les exploitations (Laborde, 2012).

Depuis le milieu des années 2000, des systèmes, dits de « monitoring », ont été développés pour assister les éleveurs pour la détection d'états physiologiques (vêlages, chaleurs) ou pathologiques (maladies hyperthermisantes, boiteries, maladies métaboliques) (Rutten et al., 2013). Ces systèmes utilisent des capteurs portés par les animaux, qui enregistrent en continu des grandeurs physiques (accélérations, sons, température, conductivité électrique) ou chimiques (pH). Ces grandeurs sont ensuite traitées automatiquement et transformées en informations relatives au comportement alimentaire (durée d'ingestion), la reproduction (alerte de chaleurs, imminence du vêlage) ou la santé (alerte mammite). Les informations élaborées fournies par les appareils de monitoring sont ensuite intégrées par l'éleveur à son système d'information, donnant accès par exemple aux performances laitières, stades de lactation, historique des troubles de santé de l'animal, informations issues de l'observation visuelle des animaux, pour prendre une décision (examiner l'animal, insémination...). Les outils de monitoring sont donc bien des outils d'aide à la détection d'évènements physiologiques ou pathologiques qui sont intégrés au système d'information de l'élevage et donnent des informations complémentaires pour la gestion de la santé et de la reproduction. Différents niveaux de développement des outils peuvent être définis en fonction du degré d'élaboration de l'information fournie aux éleveurs (Rutten et al., 2013). Cette information peut varier de l'information brute issue du capteur à une prise de décision autonome de l'outil. La finalité de ces outils est en effet bien de fournir une aide à la décision pour l'éleveur.

Les principales mesures pouvant être effectuées sur les vaches laitières sont représentées sur la figure 1-1. Certains dispositifs sont pour le moment réservés à la recherche. Une liste non exhaustive des principaux systèmes de monitoring disponibles en élevage commercial en France est présentée dans le tableau 1-1.

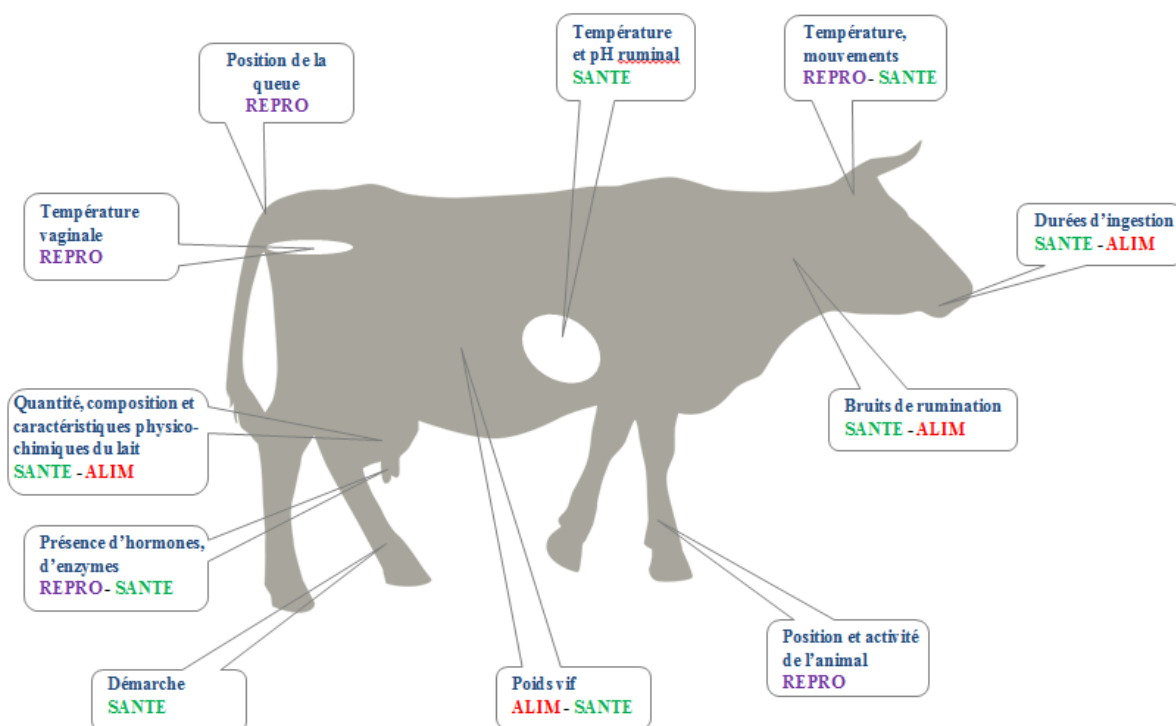


Figure 1-1 : Principales mesures pouvant être effectuées sur les vaches laitières, localisation des capteurs associés et domaines d'utilisation (d'après Idele, 2012)

La disponibilité de ces différents capteurs en élevage a fait émerger la notion d'**élevage de précision** (Wathes et al., 2008), dont il existe plusieurs définitions. Bewley (2010) définit cette notion comme « l'utilisation de technologies pour mesurer des indicateurs physiologiques, comportementaux et de production en vue d'améliorer le management et les performances des exploitations agricoles ». Eastwood et al. (2004) ont quant à eux défini l'élevage de précision comme « l'utilisation des technologies de l'information pour la mesure fine de la variabilité liée à l'animal et aux ressources dans le but d'améliorer les stratégies de management et *in fine* optimiser les performances économiques, sociales et environnementales des exploitations ».

Tableau 1-1 : Panorama des principaux outils de monitoring disponibles et leurs applications (cf annexe générale)

Système	Mesure	Applications	Coût (50 vaches)
<i>Paramètres physiologiques sur l'animal</i>			
Compteur à lait	Production laitière	Pilotage du troupeau (santé, alimentation, performance)	500 à 1500€ par poste ou intégré au robot
Thermomètre	Température ruminale ou auriculaire	Troubles infectieux	5 à 10k€
	Température vaginale	Vêlages	3 à 4 k€ (+ abonnement éventuel)
Bolus ruminal	pH ruminal	Troubles métaboliques	27 k€
<i>Paramètres biologiques sur les produits</i>			
Analyseur de lait	Concentration en cellules somatiques	Mammites	7500€ + consommables
	Conductivité		500 à 1500€ par poste ou intégré au robot
	Matière grasse, matière protéique	Troubles métaboliques	-
	Dosage de progestérone, Lactate déshydrogénase, β -hydroxy butyrate, urée	Chaleurs Mammites Troubles métaboliques	40 à 50 k€ (+50 €/vache/an)

Paramètres morphologiques			
Balance	Poids	Troubles métaboliques Gestion de l'alimentation, Performances (bovins allaitants)	5 à 10 k€
Paramètres comportementaux			
Ceinture abdominale	Contractions abdominales	Vêlages	3,5 à 4 k€
Podomètre	Activité motrice (nombre de pas)	Chaleurs	4,5 à 8 k€ (10 k€ si tous animaux équipés)
	Position debout/couché	Vêlage Boiteries Chaleurs	4,5 à 8 k€ (10 k€ si tous animaux équipés)
Accéléromètre	Activité Physique	Chaleurs Boiteries	4,5 à 8 k€ (+ abonnement éventuel)
	Temps d'ingestion Temps de rumination	Troubles de santé	
	Mouvements de la queue	Vêlages	3,5 à 4 k€
Détecteur électronique de chevauchement	Acceptation du chevauchement	Chaleurs	6,5 k€
Microphone	Temps de rumination	Troubles métaboliques	8 à 10 k€ (tous animaux équipés)
Tapis	Pression de chaque patte	Boiteries	-

La demande en outils de monitoring individuels de la part des éleveurs peut s'expliquer par plusieurs éléments. Tout d'abord, la charge de travail représentée par les besoins de détection des vêlages, des chaleurs, des troubles de santé, d'interventions et de traitements individuels augmente le temps d'astreinte auprès du troupeau, ce qui est une préoccupation majeure pour

l'éleveur (Cournot et Chauvat, 2012, Idele, 2010). En effet, les éleveurs, comme le reste de la société, aspirent à une diminution de leur temps de travail, ou du moins à une **organisation plus souple de leurs astreintes**. Cela se traduit par un développement des sociétés agricoles (GAEC et EARL), qui représentaient 28% des exploitations agricoles en 2007 contre seulement 7% en 1988, ou par le recours ponctuel à des services de remplacement de l'éleveur (Villaume et Delame, 2009). Le travail sur l'exploitation, et notamment les soins aux animaux, est alors partagé entre plusieurs personnes, ce qui peut poser des problèmes de surveillance des animaux, de transmission des informations entre éleveurs et d'observance et de suivi des traitements. Cela peut dégrader la détection et la prise en charge des troubles de santé et des chaleurs en entraînant une sous-détection ou une détection tardive de ces événements. Les outils de monitoring, en apportant aux éleveurs des informations utiles à la prise de décision, peuvent également diminuer la charge mentale représentée par la responsabilité de la détection d'évènements de santé ou de reproduction (chaleurs, vêlages, mammites...) (Hostiou et al., 2014).

Concernant spécifiquement la gestion de la santé animale, les outils de monitoring pourraient donc également être des leviers pour son amélioration, en permettant par exemple une détection plus précoce de certains troubles. La gestion sanitaire impacte directement l'économie des exploitations mais également le bien-être des éleveurs et des animaux (EFSA, 2009). L'impact économique d'une gestion suboptimale de la santé et de la reproduction en élevage est important. Il se compose des coûts directs (frais vétérinaires, coûts liés aux pratiques de prévention, pénalités, coût d'inséminations supplémentaires par exemple), mais également des pertes zootechniques engendrées par les maladies et les traitements (délais d'attente, réformes précoces pour les mammites par exemple). Ainsi certains travaux (Fourichon et al., 2002) rapportent un impact économique des troubles de santé et de reproduction dans un échantillon de troupeaux laitiers des pays de la Loire compris entre 28 et 34€/1000L, soit environ 10% du coût de production du lait, avec bien entendu une forte variabilité de ces chiffres suivant les situations sanitaires et les systèmes d'élevage considérés. La surveillance du troupeau est donc capitale pour optimiser non seulement la détection des chaleurs mais aussi la détection des troubles de santé de l'atelier laitier, première étape du processus de soin, avant tout diagnostic et traitement. Les conséquences d'une surveillance non optimale du troupeau sont que des animaux malades ne reçoivent aucun traitement (du fait de l'absence de détection) et que d'autres sont traités tardivement. Plusieurs conséquences peuvent découler de cette surveillance non optimale. Tout d'abord le bien-être animal est

dégradé puisque les animaux sont malades plus longtemps que s'ils étaient détectés et traités plus rapidement. Ainsi un rapport de l'EFSA (EFSA, 2009) souligne que l'occurrence de maladies, et en particulier des boiteries, mammites, maladies métaboliques et troubles de la reproduction, qui sont les plus fréquentes, constituent des atteintes majeures au bien-être des vaches laitières. Le rapport recommande la mise en place de mesures de détection précoce des maladies ainsi que la formation des éleveurs à la détection des maladies aux stades précoces dans le but d'améliorer le bien-être des vaches laitières. On peut penser que des systèmes de monitoring peuvent dans certains cas aider les éleveurs pour la détection précoce des maladies. D'autre part, les probabilités de guérison sont plus limitées en cas de traitement tardif, ce qui induit des **pertes de production** plus importantes (Timsit et al., 2011b), ainsi qu'un **usage non optimal des traitements médicamenteux**.

Ceci rejoint un enjeu de **santé publique** : le ministère français de l'agriculture a lancé en novembre 2011 un plan national de réduction des risques d'antibiorésistance en médecine vétérinaire dit plan Ecoantibio 2017, afin de rationaliser l'utilisation des traitements antibiotiques destinés aux animaux d'élevage et de réduire ainsi les risques de sélection de bactéries résistantes chez l'homme. Des objectifs chiffrés de réduction des volumes d'antibiotiques ont été annoncés (-25% en 5 ans). Les dispositifs d'aide à la détection des animaux malades pourraient donc contribuer à l'individualisation (traitement sélectif) et à l'optimisation des traitements médicamenteux chez les bovins (traitements curatifs ou métaphylactiques ciblés, traitements précoces moins longs et ayant plus de chances de succès). Au contraire, si ces outils détectent des animaux non malades, ils pourraient conduire à une augmentation des traitements médicamenteux.

2. Positionnement et problématique générale du travail de recherche

2.1 Objectifs et contexte du projet de thèse

Ce travail de thèse est issu de la collaboration entre l'entreprise Evolution (Rennes, France), qui a financé le travail, et l'UMR1300 BioEpAR (Nantes, France) qui en a assuré l'encadrement scientifique. Pour l'entreprise Evolution, l'objectif général de la thèse était

d'évaluer les potentialités d'un outil de monitoring dont elle possède la distribution exclusive en France : le Heatime-HR®.

2.2 Description du système de monitoring support de la thèse

L'outil support de cette thèse, Heatime-HR®, mesure en continu **l'activité générale et la rumination** des animaux équipés et fournit des alertes de chaleurs et de santé à partir de ces données. Une description détaillée de l'appareil est fournie en annexe 1.1. Brièvement, les vaches sont équipées de colliers qui portent un boîtier en plastique situé sur l'encolure. Le boîtier contient 2 capteurs : 1 accéléromètre qui mesure les mouvements de l'animal dans les 3 dimensions de l'espace et 1 microphone qui mesure les sons. Les données brutes issues des capteurs sont ensuite traitées automatiquement (filtrages de certaines fréquences pour les sons par exemple) pour donner une donnée d'activité (sans unité) et une durée de rumination (minutes) par tranche de 2 heures. Les données sont ensuite transférées périodiquement du boîtier individuel à une base informatique qui stocke et permet de visualiser les données de chaque vache équipée sous forme de courbes. L'outil générait initialement des alertes pour la détection des chaleurs. Depuis mi-2011, il applique sur ces données des algorithmes pour la détection des troubles de santé, et fournit à l'éleveur des listes de vaches détectées en chaleur ou en alerte de santé. Un algorithme détecte les hyperactivités liées aux chaleurs et les vaches peuvent apparaître sur la liste des vaches en alerte de santé en cas de baisse de rumination et/ou d'activité. Les algorithmes de détection sont la propriété de l'entreprise qui a développé le Heatime-HR® (SCR Engineers Ltd, Netanya, Israël) et ne sont donc pas connus dans le détail.

On peut donc distinguer dans l'outil Heatime-HR® 2 composantes : une première partie « capteurs » qui fournit 2 données (activité/2h, durée de rumination/2h) utilisables pour la recherche, et une seconde partie « logicielle » (interface utilisateurs et algorithmes de détection des chaleurs et des troubles de santé) qui évolue suivant les développements effectués par l'entreprise SCR Engineers et qui est mal connue et peu accessible pour la recherche.

2.3 Compétences scientifiques mobilisables autour du monitoring

Le monitoring de la santé et de la reproduction consiste, à partir de mesures effectuées sur les animaux et éventuellement d'autres informations disponibles dans le système d'information de l'élevage, à fournir des informations élaborées comme par exemple des listes de vaches à surveiller. Le passage de la mesure élémentaire effectuée sur les animaux à une information élaborée fait intervenir plusieurs disciplines.

L'électronique et les télécommunications : les mesures effectuées sur les animaux reposent sur l'utilisation de capteurs mesurant « en continu », c'est-à-dire avec une fréquence élevée au cours de la journée, une ou plusieurs grandeurs physiques (température, activité, conductivité électrique, son...) ou chimiques (pH). Les données issues du capteur sont ensuite transmises à distance, par des technologies utilisant différentes longueurs d'ondes, à des bases réceptrices de radiofréquences, à un ordinateur ou à un téléphone portable, qui permettent la consultation des données.

La physiologie et la pathologie : les données issues de capteurs remettent en cause les connaissances acquises concernant les fluctuations normales et pathologiques de grandeurs comme la température corporelle en donnant accès à des mesures « en continu » sur un plus grand nombre d'animaux qu'antérieurement. Il est dans un premier temps nécessaire de décrire les données obtenues chez des animaux apparemment sains afin de mettre à jour les référentiels des valeurs physiologiques de la grandeur mesurée. La variabilité interindividuelle des températures corporelles des bovins était, par exemple, mal connue avant les travaux portant sur les capteurs de température. Ces données permettent également de décrire avec plus de précision des phénomènes de santé, en permettant par exemple de décrire les hyperthermies (localisation dans le temps, durée) lors de broncho-pneumonies chez des jeunes bovins à l'engrais (Timsit et al., 2011a, Timsit et al., 2011b). D'autre part, certains capteurs donnent accès à de nouvelles données. Ainsi les durées de rumination ou le pH ruminal, qui étaient uniquement mesurables par le passé dans un cadre expérimental, sont désormais mesurables en élevage sur un plus grand nombre d'animaux.

La génétique peut également trouver un intérêt dans les mesures automatiques effectuées sur les animaux. En effet, si certaines des mesures faites grâce aux capteurs étaient déjà réalisées « manuellement » par le passé (production laitière, taux de matières utiles du lait, concentration en cellules somatiques du lait par exemple), d'autres comme l'activité générale,

le pH ruminal ou la rumination ne l'étaient pas. D'autre part, la fréquence des mesures était généralement beaucoup plus faible car celles-ci étaient beaucoup plus contraignantes et/ou coûteuses. Les nouvelles données pourraient permettre dans les années qui viennent le **phénotypage à haut débit**. En effet, le développement récent de la sélection génomique des animaux, dans laquelle la valeur génétique des individus peut être estimée avec une bonne précision dès la naissance grâce à une simple prise de sang, modifie en profondeur la structure des programmes de sélection génétique bovine : raccourcissement de l'intervalle de génération, diminution de la population de référence (20 000 à 30 000 individus seulement), possibilité de sélectionner des animaux pour lesquels le phénotype n'a pas été mesuré sur eux-mêmes ou des individus apparentés, et augmentation du nombre de caractères sélectionnables (Boichard et Brochard, 2012). Une des manières d'augmenter le nombre de caractères sélectionnables par des mesures sur animaux vivants en élevage est le phénotypage à haut débit qui consiste à effectuer, grâce à des capteurs équipant les animaux ou les systèmes de collecte de lait, des mesures répétées, automatisées et rapides de caractères mesurables (Hocquette et David, 2011). Ainsi plusieurs projets sont en cours concernant l'analyse de la composition fine du lait par analyse des spectres infrarouge (programme européen Phénofinlait), l'enregistrement des troubles de santé, la mesure électronique de l'état corporel (Boichard et Brochard, 2012). L'enjeu pour la filière bovins lait est de sélectionner des animaux répondant aux contraintes et attentes sociétales de demain : qualité des produits, optimisation de l'utilisation des ressources, robustesse des animaux, bien-être animal, adaptation aux changements climatiques, réduction de l'empreinte environnementale...

Les mathématiques et les statistiques : la production d'alertes à partir des mesures répétées issues des capteurs passe par un traitement mathématique des données avec application d'algorithmes, précédé ou non d'une étape de modélisation des données. L'objectif du traitement mathématique des données est de détecter des valeurs observées déviantes par rapport aux valeurs attendues en l'absence de perturbation.

Les algorithmes d'analyse des données mis en œuvre pour la détection des événements de santé ou de reproduction ont pu varier depuis des techniques simples jusqu'aux méthodes logique floue : essai de seuils empiriques (par exemple, sur l'activité et la détection de l'œstrus pour Roelofs et al. (2005), ou sur la teneur du lait en cellules somatiques et présence d'infections intramammaires, Mollenhorst et al., (2010)) ; moyennes mobiles (par exemple, conductivité du lait et présence d'infections intramammaires, (Lukas et al., 2009)) ; réseaux de neurones (par exemple Sun et al. (2010), toujours pour la conductivité du lait et les

infections intramammaires) ; méthodes de régression pour l'analyse de séries temporelles (par exemple, pour différents troubles de santé et pour la détection de l'oestrus, de Mol et al. (1999)) ; logique floue (par exemple Kamphuis et al. (2008), également pour les infections intramammaires) ; et enfin les méthodes de « data mining ». Le choix des algorithmes adéquats est difficile, d'une part à cause de la multitude de méthodes utilisables, et d'autre part parce qu'il est impossible de déterminer *a priori* quelle est la méthode la plus adaptée aux données. Les méthodes utilisées doivent cependant ne pas être sensibles à la présence de données manquantes, qui est fréquente lorsque l'on utilise des données issues de capteurs.

L'épidémiologie : l'évaluation et le paramétrage des algorithmes de détection se fait sur la base de connaissances épidémiologiques. Des indicateurs comme la sensibilité (aptitude à détecter tous les animaux malades), la spécificité (aptitude à ne détecter que les animaux malades), les valeurs prédictives positives (pourcentage d'animaux en alerte réellement malades) et négatives (pourcentage d'animaux non détectés réellement non malades) des détections permettent de paramétrer au mieux les algorithmes en fonction des objectifs poursuivis, et notamment des coûts des détections à tort (faux positifs ou faux négatifs). En effet, suivant les cas il peut être préféré un appareil à sensibilité élevée détectant bien les animaux malades ou un appareil spécifique qui détecte peu d'animaux non malades. Un paramétrage adéquat de la sensibilité et de la spécificité est important pour un usage réel de l'outil de détection par l'éleveur. En effet, les performances de détection de l'outil impactent (i) directement la confiance que les éleveurs peuvent accorder aux informations fournies par l'appareil et donc l'utilisation réelle de l'appareil en élevage. En présence d'un outil trop peu spécifique, beaucoup d'animaux en alerte seront considérés sains après examen par l'éleveur et auront été examinés « pour rien », ce qui est coûteux en temps. En présence d'un outil trop peu sensible, des animaux nécessitant des soins ne seront pas détectés par l'appareil, ce qui peut entraîner un retard à la détection voir une non détection par l'éleveur. (ii) indirectement l'impact économique des troubles de santé dans les exploitations équipées, le bien-être des animaux et l'usage des traitements médicamenteux en élevage. Les dispositifs d'enquête épidémiologique sont également à questionner lors de l'évaluation de dispositifs de monitoring, en raison de leurs caractéristiques (mesures en continu, détection potentielle de plusieurs maladies à la fois, prise en compte des facteurs pouvant impacter la mesure).

L'économie et la sociologie : les outils de monitoring peuvent impacter l'économie des exploitations car ils peuvent représenter des investissements conséquents et leur impact potentiel sur les performances technico-économiques des exploitations est rarement évalué.

Comme toute nouvelle technologie, ils modifient profondément le travail des agriculteurs et les outils de la sociologie peuvent aider à comprendre quels usages sont faits de ces outils et quels sont leurs apports ou les freins à leur utilisation.

Le monitoring se situe donc au carrefour de différentes sciences et un travail de thèse ne peut prétendre aborder l'ensemble des champs scientifiques d'intérêt lors de l'évaluation d'un outil de monitoring. Nous avons donc restreint notre analyse à certains domaines de connaissance.

2.4 Positionnement du travail de recherche

Parmi les domaines de connaissance rattachés au monitoring à l'aide de cet outil, nous avons privilégié 2 aspects essentiels :

Le phénotypage : les données brutes (activité/2h, durée de rumination/2h) fournies par l'appareil de monitoring pouvaient présenter un intérêt pour le distributeur de Heatime-HR® en France, qui est une entreprise de sélection bovine. Il s'agissait d'évaluer l'intérêt des phénotypes mesurés par l'outil dans une optique de valorisation ultérieure en sélection bovine.

L'épidémiologie : nous avons vu dans la partie précédente l'importance de l'évaluation de la valeur informative des alertes de santé ou de reproduction pour une utilisation réelle par les éleveurs. L'utilisation des outils épidémiologiques dans cette thèse a également été favorisée par l'encadrement scientifique du travail au sein d'une unité de recherche en épidémiologie (UMR 1300 BioEpAR). Evaluer les algorithmes commerciaux présentait plusieurs difficultés. Tout d'abord de nouveaux algorithmes pouvaient être produits par le fabricant à tout moment, rendant caduque l'évaluation en cours. Ceci est d'autant plus problématique que, les troubles de santé chez les vaches laitières étant rares, un suivi relativement long des élevages est nécessaire. Ainsi en 2009, lors du début de la collecte des données utilisées dans cette thèse, l'appareil Heatime-HR® fournissait uniquement des données brutes de rumination et d'activité et des alertes de chaleur. Les alertes de santé ne sont apparues qu'en 2011, et nous n'avons pas eu connaissance dans le détail des modifications survenues dans les algorithmes depuis 2011. D'autre part travailler directement avec les alertes fournies par l'appareil ne permettait pas d'explorer différents réglages de sensibilité ou de spécificité des détections. Pour toutes ces raisons, nous avons préféré développer et évaluer nos propres algorithmes de détection des troubles de santé au cours de cette thèse.

Ces 2 domaines de connaissance ont été privilégiés sur la base de leur importance pour l'entreprise Evolution, distributrice de Heatime-HR® en France, qui souhaitait évaluer le potentiel de l'outil pour :

- Le phénotypage et la sélection animale
- La détection des troubles de santé

Les compétences en traitement du signal et en développement d'algorithmes n'étant pas le cœur de métier de l'UMR1300 BioEpAR, cette partie n'a pas constitué la partie centrale de la thèse. La partie **mathématique et production d'alertes** a été faite en partenariat avec l'équipe DREAM (René Quiniou, INRIA, Rennes, France).

Dans la suite de ce manuscrit, nous appellerons **valeur informative de l'activité et des durées de rumination (ou des alertes produites à partir de ces données)** l'ensemble des connaissances permettant d'utiliser ou d'interpréter les données d'activité ou de rumination (ou les alertes associées). Pour les données brutes de rumination, il pourra s'agir d'une méthode permettant éventuellement de mesurer un phénotype d'intérêt à partir de la donnée issue du capteur. Pour les alertes de santé ou de reproduction, la valeur informative des alertes sera l'ensemble des statistiques relatives à l'évaluation de la performance des alertes (sensibilité, spécificité...).

2.5 Problématique générale

L'activité générale et les durées de rumination des vaches laitières sont des données mesurables depuis assez peu de temps en ferme commerciale. Par conséquent, la valeur informative de ces données, ou d'alertes pouvant être générées à partir de ces données, est mal connue. La problématique générale de cette thèse **est de déterminer quelle est la valeur informative de l'activité générale et de la durée de rumination mesurée sur les vaches laitières, ou d'alertes pouvant être générées à partir de ces données.**

Nos principales hypothèses de travail étaient que les durées de rumination puissent être informatives pour prédire les quantités ingérées ou l'efficacité alimentaire des animaux, et que

l'activité générale et les durées de rumination puissent être informatives pour la détection de troubles de santé.

3. Etat des connaissances et questions de recherche sur la valeur informative de l'activité générale et des durées de rumination

Comme toute mesure, les mesures réalisées par Heatime-HR® sont imparfaites en raison des capteurs eux-mêmes ou des algorithmes qui traduisent les grandeurs physiques mesurées (mouvement et son) en données élaborées que sont l'activité et les durées de rumination.

3.1 Qualité de la mesure de l'activité par Heatime-HR®

Dans une étude récente, Elischer et al. (2013) ont tenté de corréler les données d'activité mesurées par les colliers Lely, qui utilisent les mêmes capteurs (HR-Tags, accéléromètre et microphone situés à l'encolure) que le système Heatime-HR®, avec des comportements relevés par observation visuelle ou par un podomètre (IceQube, IceRobotics, Edinburg, UK). L'essai portait sur 15 vaches Holstein en robot au pâturage, observées pendant 7 périodes de 2 heures chacune. Des corrélations modérées étaient trouvées entre l'activité mesurée par l'accéléromètre et les observations de marche ($R^2=0.37$, $n=105$), station debout ($R^2=0.21$), station couchée ($R^2=0.32$) et démarche active, c'est-à-dire la combinaison de la marche ou de la station debout ($R^2=0.27$). Les corrélations étaient également modérées entre l'activité mesurée par l'accéléromètre d'encolure et le nombre de pas ou l'activité mesurés par le podomètre. Cependant, ce résultat n'est pas vraiment étonnant puisqu'une valeur d'activité, sans unité, issue de l'accéléromètre de cou et ne prétendant pas correspondre à un comportement biologique, était corrélée à des activités biologiques, exprimées en pourcentage de temps d'expression sur la période d'observation de 2 heures. De plus, l'accéléromètre d'encolure a été conçu pour la détection d'une hyperactivité liée aux chaleurs et non pour l'enregistrement de comportements biologiques.

3.2 Qualité de la mesure des durées de rumination par Heatime-HR®

Le tableau 1-2 synthétise les résultats obtenus par le système Heatime-HR® lors d'études de validation de la mesure des durées de rumination.

Tableau 1-2. : synthèse des études de validation des durées de rumination mesurées par le système Heatime-HR®

Etude	Méthode de référence ¹	Types d'animaux	Logement	Nombre d'animaux	Nombre de périodes de 2 h ou de jours	Type de ration ²	R ²
Schirmann et al., 2009	OV	Holstein adultes	Box	15 12	n=15x2h n=36x2h	RC RC	0,93 0,86
Byskov et al., 2014	Licol mesurant la rumination	Génisses Jersiaises	Attache	4	n=1005x2h n=37 j	4 rations à base d'ensilage d'herbe ou foin	0,83 0,62
Elischer et al., 2013	OV	Holstein adultes	Stabulation libre	15	n= 63x2h	P	0,42
Burfeind et al., 2011	OV	Génisses 9 mois	Box	5	n=15x2h	RC	0,77
Goldhawk et al., 2013	OV (vidéo en continu)	Génisses Angus-Hereford	Attache	7	n=119x2h	Riche en fourrage (ensilage orge) ou en céréales (orge)	0,006 à 0,21
	OV	Broutards Angus-Hereford	Case collective	6	n=44x2h	Riche en fourrage (ensilage orge)	0,17

¹OV= Observation visuelle

²RC=Ration Complète, P=Pâturage principalement

Schirmann et al. (2009) ont comparé les durées de rumination obtenues par Heatime-HR® et par observation visuelle pour 51 périodes de 2h sur 27 vaches Holstein. Les coefficients de corrélation obtenus au cours des 2 essais de l'étude étaient élevés (respectivement $R^2=0,93$, $n=15$ et $R^2=0,86$, $n=36$). Le biais entre observation visuelle et Heatime-HR® ne dépendait pas de la valeur de la durée de rumination sur la période considérée. Les corrélations obtenues sont comparables à celles obtenues pour d'autres dispositifs de mesure des durées de mastication (Beauchemin et al., 1989). Cependant, la validation a été effectuée uniquement dans un élevage, sur un nombre limité de périodes de 2h, chez des vaches mises en cases individuelles et nourries en ration complète. La qualité de la mesure qui serait obtenue en conditions de terrain, pour des vaches en troupeau recevant une ration conservée ou pâturant suivant la période de l'année considérée, ou pour des vaches recevant une ration alimentaire différente des rations utilisées lors de la validation est inconnue.

Byskov et al. (2014) ont validé la mesure des durées de rumination chez 4 génisses Jersiaises à l'attache, recevant différents fourrages, avec ou sans restriction alimentaire. Les données étaient fournies par des colliers vendus avec les robots de traite Lely, mais qui utilisent la technologie SCR (capteurs HR-Tags) et fournissent les mêmes données que Heatime-HR®. La méthode de référence était un collier enregistrant automatiquement l'activité masticatoire des animaux, ce collier ayant fait l'objet d'une validation sur quelques animaux précédemment. Des coefficients de corrélations élevés étaient obtenus entre les 2 méthodes de mesure ($R^2=0,83$, $n=1005$ à l'échelle de périodes de 2 heures, $R^2=0,62$, $n=37$ à l'échelle de la journée). La nature des fourrages distribués, la période de la journée, et l'existence d'une restriction alimentaire influençaient la qualité de la mesure des durées de rumination. Les hypothèses avancées étaient qu'une teneur en NDF (Neutral Detergent Fiber) élevée des fourrages pouvait augmenter les sons de rumination, que des activités comme l'ingestion, la buvée ou la station couchée pouvaient interférer avec la mesure de la rumination. La mesure pourrait également varier en fonction de caractéristiques individuelles, puisque la durée de rumination mesurée était systématiquement surestimée (entre 39 et 139 min/j) pour une des 4 génisses de l'étude.

Elischer et al. (2013) ont également tenté de corrélérer les durées de rumination mesurées par les colliers Lely avec celles mesurées par observation visuelle. L'essai portait sur 15 vaches Holstein en robot au pâturage, observées pendant 7 périodes de 2 heures chacune. Le coefficient de corrélation entre les durées de rumination enregistrées par le collier et celles

enregistrées par les observateurs était de 0,42 (n=63), avec parfois plus de 30 minutes d'écart entre les 2 mesures sur une période de 2 heures d'observation.

Une étude de validation faite chez de jeunes veaux et génisses a été publiée (Burfeind et al., 2011). Les durées de rumination étaient mesurées par Heatime-HR® et observation visuelle chez des animaux âgés de 4 semaines à 9 mois. Pour les animaux sevrés, des coefficients de corrélations élevés entre les 2 méthodes de mesure étaient obtenus uniquement sur 1 lot de 5 animaux de 9 mois ($R^2=0,77$, n=15 périodes de 2h). Les auteurs concluaient que la mesure automatisée des durées de rumination par HR-Tag était suffisamment précise après l'âge de 9 mois.

Goldhawk et al. (2013) ont évalué la qualité de la mesure des durées de rumination par Heatime-HR® chez des génisses et des taurillons recevant différentes rations. Dans la première partie de l'étude, les durées de rumination de 7 génisses Angus ou Angus-Hereford de première année (n=119 périodes de 2h) logées en étable entravée étaient comparées aux durées de rumination obtenues par observation visuelle. Dans la seconde partie de l'étude, les mesures étaient effectuées sur 6 broutards de la même race, en liberté dans 1 enclos typique de feedlot (n=44 périodes de 2h). En moyenne, les durées de rumination étaient de 16,7 min/2h avec Heatime-HR® contre 26,6 min/2h par observation visuelle, soit une sous-estimation des durées de rumination de 9,8 min par période de 2 h par Heatime-HR®. Les coefficients de corrélation obtenus au cours de l'essai variaient entre 0,006 et 0,212 suivant le lot considéré, avec un coefficient de corrélation global de 0,168. Il n'y avait pas de relation entre le biais de mesure par Heatime-HR® et la position de la tête des animaux, malgré que les auteurs aient observé que les capteurs de rumination ne restaient pas localisés à la place idéale sur le cou de l'animal lorsque ceux-ci baissaient la tête. Les auteurs émettaient l'hypothèse que la mauvaise performance de Heatime-HR® pouvait être liée à des facteurs alimentaires, les rations utilisées dans leur étude étant moins riches en fourrages que les rations des 2 autres études de validation, ainsi que des facteurs liés à l'âge ou aux différences physiques entre vaches laitières et animaux allaitants (fanon, musculature du cou, épaisseur de la peau).

Ainsi, même si le nombre d'études de validation est limité, il semble que Heatime-HR® ait une précision assez bonne pour la mesure des durées de rumination chez les vaches laitières adultes. La qualité de la mesure en valeur absolue peut être importante pour le phénotypage par exemple, pour lequel on veut une mesure standardisée sur les animaux. Au contraire, si

l'on veut détecter des variations relatives de durées de rumination d'un animal, par exemple pour la détection des troubles de santé, l'erreur absolue de la mesure est moins importante que la répétabilité de l'erreur. En effet, une erreur systématique ne posera pas de problème si l'on cherche à détecter des variations relatives de la durée de rumination. Aucune étude ne permet aujourd'hui de juger de la répétabilité de l'erreur de mesure des durées de rumination intra-vache.

La méthode de référence utilisée varie également suivant les études de validation. On peut considérer que le Gold Standard pour mesurer les durées de rumination est l'observation visuelle. En effet, les appareils automatiques comme celui utilisé par Byskov et al. (2014) commettent une erreur de mesure qui peut parfois être non négligeable (mal quantifiée pour le dispositif utilisé par Byskov et al. (2014)) et peuvent également poser des problèmes de fiabilité liés à la complexité technique des dispositifs.

3.3 Intérêt de l'activité générale et des durées de rumination en termes de phénotypage

L'activité générale et les durées de rumination pourraient présenter un intérêt pour évaluer respectivement le niveau d'expression des chaleurs et les quantités ingérées ou l'efficacité alimentaire des vaches laitières. En effet, l'activité tout d'abord, est mesurée de façon standardisée par l'accéléromètre sur toutes les vaches équipées du collier Heatime-HR®. On sait que les chaleurs sont de plus en plus discrètes et de plus en plus courtes chez les vaches laitières (Lopez et al., 2004, Palmer et al., 2010), ce qui peut poser un problème pour leur détection dans certains élevages. Les troubles de la cyclicité sont également des troubles majeurs de la reproduction des vaches laitières (Shrestha et al., 2004). On peut imaginer qu'il soit possible grâce à ce système de mesurer finement un phénotype de niveau d'expression des chaleurs (intensité des chaleurs, niveau du pic d'activité, durée de suractivité) ou un phénotype lié à la cyclicité des chaleurs. Si ces phénotypes possèdent un déterminisme génétique suffisant, il serait alors possible d'indexer les animaux sur leur niveau d'expression des chaleurs et de sélectionner les animaux présentant un index élevé pour la détectabilité de l'œstrus.

Plusieurs facteurs liés à la composition de la ration, comme la qualité d'un fourrage ou sa teneur en fibres (Adin et al., 2009), peuvent faire varier les durées de rumination des vaches. Si l'impact de différentes rations sur les durées de rumination a été largement étudié, **la relation entre les durées de rumination et les quantités ingérées a été peu explorée.**

Cependant, il est fondamental d'explorer l'existence d'une telle relation. En effet, si une relation existait, la mesure des durées de rumination pourrait être un moyen de mesurer indirectement l'ingestion des animaux. Or l'estimation de l'ingestion individuelle est capitale à la fois pour la formulation optimale de rations alimentaires et éventuellement pour le calcul de l'efficacité alimentaire des animaux. **Les applications pour le calcul de l'efficacité alimentaire individuelle pourraient être très importantes pour les entreprises de sélection de vaches laitières (Gonzalez-Recio et al., 2014, Pryce et al., 2014).** En effet, le calcul de l'efficacité alimentaire des vaches laitières est désormais possible à partir de l'ingestion individuelle, du poids vif, de la quantité et de la composition du lait produit. Parmi ces mesures, les pesées automatisées ou l'imagerie 3D permettent d'obtenir le poids vif ou la note d'état corporelle de façon satisfaisante, et les compteurs à lait associés à des analyses des taux permettent d'obtenir la quantité et la composition du lait produit. **La mesure de l'ingestion individuelle est aujourd'hui la mesure la plus difficile à réaliser pour calculer l'efficacité alimentaire des animaux, premier pas vers une sélection sur ce phénotype.**

Seule une étude de Schirmann et al. (2012) s'est intéressée à la relation entre durée de rumination et quantité de matière sèche ingérée sur 42 vaches Holstein durant le tarissement. L'hypothèse des auteurs était que les durées de rumination augmenteraient avec les quantités ingérées, puisqu'une plus grande quantité de nourriture devrait être régurgitée et mastiquée au cours de la rumination. En analysant conjointement toutes les vaches, il n'y avait pas de relation entre durées de rumination et quantités ingérées à l'échelle de la journée. Intra-vache et à l'échelle de la journée, il y avait une faible corrélation négative entre les 2 variables. **Même si les résultats de cette étude ne sont pas en faveur de l'existence d'une relation entre durées de rumination et quantités ingérées chez les vaches tarées, il est fondamental de confirmer ces résultats et d'évaluer la relation entre durées de rumination et quantités ingérées au cours de la lactation.**

L'activité générale et les durées de rumination pouvant être impactés par une grande variété de troubles de santé (Radostits et al., 2007), un outil de monitoring comme Heatime-HR® pourrait également permettre de mesurer des **phénotypes d'intérêt en rapport avec la santé des animaux, la résistance ou la résilience aux maladies.** Les données pourraient en effet permettre de mesurer en continu et de manière standardisée des états de santé (fréquence ou durée d'alertes de santé par exemple). Des bases génétiques à la résistance à certaines maladies pouvant être suspectées, par exemple pour les mammites cliniques qui sont

désormais l’objet d’une indexation en France, une amélioration du niveau sanitaire des troupeaux par sélection pourrait être envisagée.

3.4 Valeur informative et intérêt pour la détection des chaleurs

Plusieurs études ont évalué la valeur informative des alertes de chaleurs fournies par Heatime-HR® ou par Heatime®, qui correspond au produit développé initialement par la société SCR pour la détection des chaleurs spécifiquement, et qui utilise une technologie très proche de celle utilisée par Heatime-HR® (tableau 1-3). Les 2 systèmes utilisent le même accéléromètre pour la mesure de l’activité. Cependant Heatime-HR® contient en plus un système de mesure de la rumination, ce qui fait que les boîtiers contenant les capteurs sont différents dans les 2 systèmes. Par ailleurs, les parties logicielles de ces 2 produits sont également différentes, même si la détection des chaleurs par Heatime-HR® est pour le moment uniquement basée sur la détection d’épisodes d’hyperactivité. Des performances de détections différentes peuvent donc être attendues pour les 2 outils.

Tableau 1-3: synthèse des valeurs informatives obtenues pour la détection des chaleurs à l’aide des systèmes Heatime ou Heatime-HR®

Etude	Appareil	Méthode de référence ¹	Type d'élevage	Nombre de troupeaux (d'animaux)	Synchronisation des chaleurs	Se (%)	Sp (%)	VPP (%)	VPN (%)
Aungier et al., 2012	Heatime	PL	Pâturage	1 (89 Holstein)	Non	72	-	78	-
Holman et al., 2011	Heatime	PL	Stabulation libre	1 (67 Holstein)	Oui en partie	59	-	94	-
Chanvallo n et al., 2014	Heatime-HR®	PL	Stabulation libre	1 (63 Holstein)	Non	62	-	84	-
Kamphuis et al., 2012	Heatime	PL	Pâturage	1 (320 Holstein xJersiaise)	Non	62	>99	64 à 77	-
	Heatime-HR®	PL	Pâturage	1 (315 Holstein xJersiaise)	Non	77	>99	72 à 82	-

¹PL= suivi de la concentration en progestérogène dans le lait

Aungier et al. (2012), dans une étude menée sur 89 vaches Holstein au pâturage, ont trouvé que Heatime détectait 72% des ovulations (définies à partir de dosages bi-hebdomadaires de progestérone dans le lait) et que 32% des épisodes de suractivité correspondaient à des niveaux de progestérone élevés (faux positifs). L'absence d'infection utérine, un rang d'ovulation supérieur ou égal à 2 (postérieur à la reprise de cyclicité), une augmentation de la note d'état corporel et une baisse de la production laitière étaient associés à une augmentation du risque de détection de chaleur par l'outil (OR>1). On peut ici remarquer que l'évaluation d'un appareil de détection d'un comportement lié à l'œstrus (hyperactivité) se fait par rapport à une méthode de référence qui permet de mettre en évidence les ovulations (progestérone du lait), c'est-à-dire pas exactement ce que détecte l'appareil. On sait par exemple que jusqu'à 30% des ovulations peuvent ne pas être accompagnées de comportements d'œstrus chez la vache (Ranasinghe et al., 2010, Shipka, 2000).

En conditions d'élevage, Holman et al. (2011) ont évalué les performances de détection de différentes méthodes de détection des ovulations chez 67 vaches laitières suivies pendant 6 mois (189 chaleurs). La sensibilité de tous les dispositifs de détection, dont le Heatime, la détection par observation visuelle et un podomètre, était équivalente, située autour de 60%. Par rapport aux autres méthodes d'aide à la détection, le Heatime générait le moins de fausses alertes et obtenait donc une valeur prédictive positive significativement plus élevée que les autres dispositifs (93% vs 74% pour le podomètre). La méthode de détection qui obtenait les plus fortes sensibilités (75%) et VPP (92%) était la combinaison Heatime-détection par observation visuelle. Or cette combinaison est généralement celle qui est utilisée de manière intuitive en élevage, l'éleveur se servant de l'activimètre soit pour conforter son observation soit pour renforcer sa surveillance sur un animal détecté par l'appareil.

Deux études ont évalué l'intérêt de Heatime pour la détection des ovulations dans le cadre d'élevages utilisant des protocoles hormonaux de déclenchement des chaleurs. Fricke et al. (2014) ont comparé les résultats obtenus lorsque les protocoles hormonaux n'étaient utilisés qu'en dernier recours, après une période d'insémination sur chaleurs détectées par le Heatime, aux résultats obtenus avec protocole hormonal systématique. Les résultats montrent que le système Heatime permettait une mise à la reproduction plus rapide des vaches mais que le ratio nombre de gestations par IA était meilleur chez les vaches ayant été inséminées après un protocole de synchronisation seul. Neves et al. (2012) ont comparé, dans 3 élevages commerciaux, les performances de reproduction de vaches inséminées soit suite à un traitement hormonal de déclenchement des chaleurs soit suite à une détection par le Heatime.

Cette étude démontre que l'intervalle vêlage-insémination fécondante est, suivant le troupeau considéré, soit identique dans les 2 systèmes de mises à la reproduction (1 troupeau sur 3), soit légèrement meilleur pour le système avec détection par le Heatime (2 troupeaux sur 3).

Chanvallon et al. (2014) ont également évalué les performances de l'appareil Heatime pour la détection des ovulations chez 63 vaches Holstein. La sensibilité trouvée était de 63% et la valeur prédictive positive des alertes était de 84%.

Une seule étude a comparé les performances de détection des ovulations obtenues par le système Heatime et par le système Heatime-HR®. Kamphuis et al. (2012) ont évalué les 2 outils dans un très grand troupeau de vaches Holstein x Jersiaises, au pâturage, avec des vêlages très groupés. Le gold standard utilisé était le dosage de la progestérone dans le lait. Suivant la durée de la fenêtre de temps prise pour la validation d'une détection (entre 24h et 72h) et le type de détecteur utilisé (Heatime ou Heatime-HR®), la sensibilité des détections variait de 52 à 77% pour une spécificité supérieure ou égale à 99%. Les VPP obtenues étaient comprises entre 63 et 82%. **On remarque ici que le calcul des valeurs informatives nécessite généralement de définir des fenêtres de temps dans lesquelles on considère que la détection observée est bien attribuable à l'évènement que l'on souhaite détecter (ici l'ovulation). La taille de ces fenêtres temporelles peut varier suivant les études ou différentes fenêtres peuvent être testées dans une même étude. Le choix d'une taille de fenêtre peut reposer sur des connaissances en physiologie (par exemple moment de l'oestrus par rapport à l'ovulation).** Dans l'étude de Kamphuis et al. (2012), la sensibilité du système Heatime-HR® était significativement plus élevée que celle du système Heatime (+15% environ) et la VPP obtenue avec Heatime-HR® était plus élevée que celle du Heatime avec une fenêtre de validation de 24h. Exceptée une possible utilisation de capteurs Heatime défectueux, aucune hypothèse n'était avancée par les auteurs pour expliquer les moins bonnes performances de Heatime. La méthode de détection des chaleurs utilisée précédemment dans la ferme, qui prenait en compte à la fois l'observation du comportement des animaux, mais également l'effacement de peinture déposée sur la queue et l'observation de patches de détection du chevauchement, était dans cette étude à la fois plus sensible et plus spécifique que les accéléromètres (Se=91%, Sp=99.8%, VPP=94.5%).

Les performances de détection des chaleurs par Heatime-HR® ou Heatime, qui utilise une technologie très proche de Heatime-HR®, ont été évaluées dans différents systèmes

d'élevage. Les performances sont à peu près identiques aux performances de détection des éleveurs par observation visuelle en élevage commercial (At-Taras et Spahr, 2001).

Par ailleurs, le tableau 1-2 permet de remarquer que **la spécificité et la valeur prédictive négative des alertes sont rarement évaluées**. Ces 2 valeurs informatives sont pourtant importantes pour quantifier le risque de détecter des animaux n'ayant pas ovulé. **Ceci est lié à une difficulté méthodologique : il est en effet difficile de définir ce que serait un évènement vrai négatif (absence de détection et absence d'ovulation) et de circonscrire un tel évènement dans le temps**. Par exemple, lors des périodes de concentration en progestérone dans le lait élevée (absence d'ovulation), des détections d'hyperactivité peuvent survenir à des intervalles très variables dans le temps ce qui pose problème pour comptabiliser les évènements vrais négatifs. Une des solutions peut être de définir un évènement vrai négatif comme une journée sans détection d'ovulation et sans alerte d'hyperactivité (Kamphuis et al., 2012).

Le produit Heatime comprend déjà **un algorithme de détection d'hypoactivité** d'un animal, présenté dans les argumentaires commerciaux comme pouvant détecter un certain nombre de troubles de santé affectant l'état général des animaux. Cependant, la véritable différenciation commerciale entre les systèmes de détection de chaleurs et Heatime-HR® se fait sur la fourniture par le système Heatime-HR®, en plus des vaches en chaleur, de listes de vaches susceptibles de présenter des troubles de santé. Ces listes sont générées à partir d'algorithmes détectant les baisses de rumination et/ou d'activité.

3.5 Valeur informative et intérêt pour la détection des troubles de santé

3.5.1 Signification clinique des baisses d'activité générale et durée de rumination

L'activité générale et les durées de rumination peuvent être impactées par différents troubles de santé. Ainsi, une diminution des durées de rumination importante peut être observée lors de dépression du système nerveux central, douleur, réticulo-péritonite traumatique, impaction du rumen, emphysème chronique, acidose aiguë, hypocalcémie, mammite colibacillaire, infection pulmonaire (Radostits et al., 2007) etc.... Plusieurs études ont mis en évidence des chutes de durées de rumination lors de maladies comme les mammites (Fogsgaard et al.,

2012) ou l'acétonémie (DeVries et al., 2009) ou des chutes d'activité lors de boiteries (Kamphuis et al., 2013b, Rutten et al., 2013). Même si activité générale et durées de rumination sont *a priori* plutôt des signes cliniques à relier à l'état général de l'animal, qui vont être impactés lors des troubles cliniques majeurs, aucune étude ne s'est intéressée à la relation entre l'intensité des troubles cliniques et les répercussions sur l'activité générale et les durées de rumination. Cela s'explique par la difficulté de mesure de ces grandeurs, qui n'étaient pas mesurables en ferme commerciale jusqu'à récemment. En effet, les études citées ci-dessus utilisaient des méthodes de mesure autres que le Heatime-HR®.

3.5.2 Peu d'études basées sur les mesures du Heatime-HR®

Plusieurs études ont décrit des associations entre des modifications des durées de rumination mesurées par Heatime-HR® et la survenue de troubles de santé. Ainsi Soriani et al. (2012) ont trouvé dans 1 troupeau que les vaches présentant des durées de rumination plus faibles avant vêlage présentaient des valeurs significativement plus élevées de certaines molécules (acides gras non estérifiés et beta-hydroxybutyrate) associées à des maladies métaboliques, ainsi qu'une incidence légèrement plus élevée de troubles de santé avant 35 jours post-partum. Cependant, ce dernier résultat portait sur un très faible nombre de troubles (10). Fitzpatrick et al. (2013) ont quant à eux mesuré à l'aide de Heatime-HR® les durées de rumination obtenues après induction de mammites cliniques sur 24 vaches, par injection dans 1 quartier de lipopolysaccharide de *E. Coli*, un lot de vaches recevant un traitement anti-inflammatoire au moment du challenge et l'autre un placebo. A l'échelle de la journée, l'injection de lipopolysaccharide n'avait pas d'effet sur les durées de rumination. Cependant, la cinétique des durées de rumination au cours de la journée de challenge était différente dans les 2 lots. Les durées de rumination diminuaient davantage dans les 8 heures suivant le challenge pour le lot témoin par rapport au lot recevant un traitement anti-inflammatoire. Les durées de rumination étaient ensuite plus élevées dans le lot témoin pour le reste de la journée. Miguel-Pacheco et al. (2014) ont quant à eux montré en conditions d'élevage une diminution significative des durées de rumination dans les 2 jours suivant le diagnostic de boiterie. Cependant l'effet estimé était très faible. D'autres travaux montrent une baisse de rumination mesurable au moment des mammites (Bar et Solomon, 2010), mais également en début de lactation chez des vaches présentant des maladies telles que l'acétonémie, des métrites ou des

rétections placentaires (Soriani et al., 2013). Cependant ces travaux ont uniquement fait l'objet de présentations en congrès et les matériels et méthodes utilisés sont relativement peu détaillés.

A notre connaissance, 2 études seulement ont évalué la valeur informative de détections générées à partir des données Heatime-HR® pour la détection des troubles de santé. La première concernait les boiteries. En effet, Van Hertem et al. (2013) ont évalué la sensibilité et la spécificité d'un modèle de détection combinant des données issues du capteur Heatime-HR® (données d'activité et de rumination) avec des variables liées à la production laitière sur un jeu de données comportant 44 épisodes de boiterie et 74 épisodes témoin sans boiterie. Les boiteries incluses étaient des boiteries d'assez forte intensité traitées par le vétérinaire de l'élevage, avec exclusion des vaches ayant présenté des troubles de santé avant l'épisode de boiterie. Le modèle comprenait 4 variables reliées à la production laitière et 3 reliées à l'activité générale ou aux durées de rumination des animaux. La sensibilité des alertes pour la détection des boiteries était de 89% et la spécificité était de 85%. Il faut cependant souligner que le modèle évalué avait été développé spécifiquement pour la détection des boiteries, et n'est pas du tout comparable aux algorithmes commerciaux fournissant les alertes Heatime-HR®. De plus, les variables les plus significatives dans le modèle étaient les variables liées à la production laitière et non pas les variables provenant du Heatime-HR®. **Par ailleurs, la précocité de la détection par le modèle n'était pas évaluée alors que la fenêtre de détection allait de 3 semaines avant à 3 semaines après la détection par les éleveurs, alors que des détections postérieures à la détection par les éleveurs présentent peu d'intérêt.**

Des résultats préliminaires d'un modèle de détection de l'acétonémie à partir des données d'activité, de rumination et de production laitière ont également fait l'objet d'une présentation en congrès (Steensels et al., 2012). Cependant la valeur ajoutée des données Heatime-HR® n'était pas vraiment évaluable puisque le modèle incluait également des données liées à la production laitière. De plus, à notre connaissance ce modèle n'a pas fait l'objet d'une publication détaillée en revue à comité de lecture.

Ainsi, alors que Heatime-HR® est présenté dans les argumentaires commerciaux comme susceptible de détecter une large gamme de troubles de santé, car combinant des mesures continues de l'activité et de la rumination des vaches laitières, très peu d'études sont

disponibles concernant le type de troubles de santé détectables au quotidien par l’outil, la valeur informative et la précocité des détections.

Il serait donc intéressant d’évaluer quels sont les types de troubles de santé détectables à l’aide de l’activité générale et des durées de rumination mesurées par Heatime-HR® et quelle est la valeur informative des détections pour les principaux troubles de santé rencontrés en élevage.

Cependant, plusieurs questions méthodologiques se posent lorsque l’on cherche à évaluer la valeur informative d’alertes de santé produites à partir de mesures faites en continu sur des animaux.

3.5.3 Points de vigilance méthodologiques pour évaluer la valeur informative d’alertes générées à partir de données issues de capteurs

3.5.3.1 Temporalité des évènements et attribution des alertes aux troubles de santé

Les informations fournies par les systèmes de monitoring sont quasiment continues, contrairement aux résultats de la méthode de référence, qui sont généralement des résultats ponctuels (par exemple examen clinique), éventuellement répétés dans le temps (par exemple comptage des cellules somatiques). Or, pour les mammites par exemple, il peut exister un décalage de plusieurs heures voir de plusieurs jours entre la modification du signal fournie par l’outil de monitoring mesurant la conductivité par exemple et la détection de la maladie par la méthode de référence, par exemple l’observation visuelle du lait. De plus, ce décalage temporel peut être variable d’une maladie à l’autre ou d’un épisode clinique à un autre pour une même maladie, ce qui rend complexe l’évaluation des alarmes fournies par un modèle de détection basé sur des données continues. **Le calcul de la valeur informative des alertes de santé peut alors se faire par la délimitation de fenêtres temporelles autour des alertes de santé** (Hogeveen et al., 2010, Kamphuis et al., 2013a). La fenêtre temporelle est alors l’unité statistique d’analyse. Si un trouble de santé est présent dans une fenêtre temporelle, celle-ci est comptabilisée comme un vrai positif, sinon elle est considérée comme un faux positif. Si aucune alerte n’est associée à un trouble de santé, ce trouble est considéré comme un faux négatif (figure 1-2).

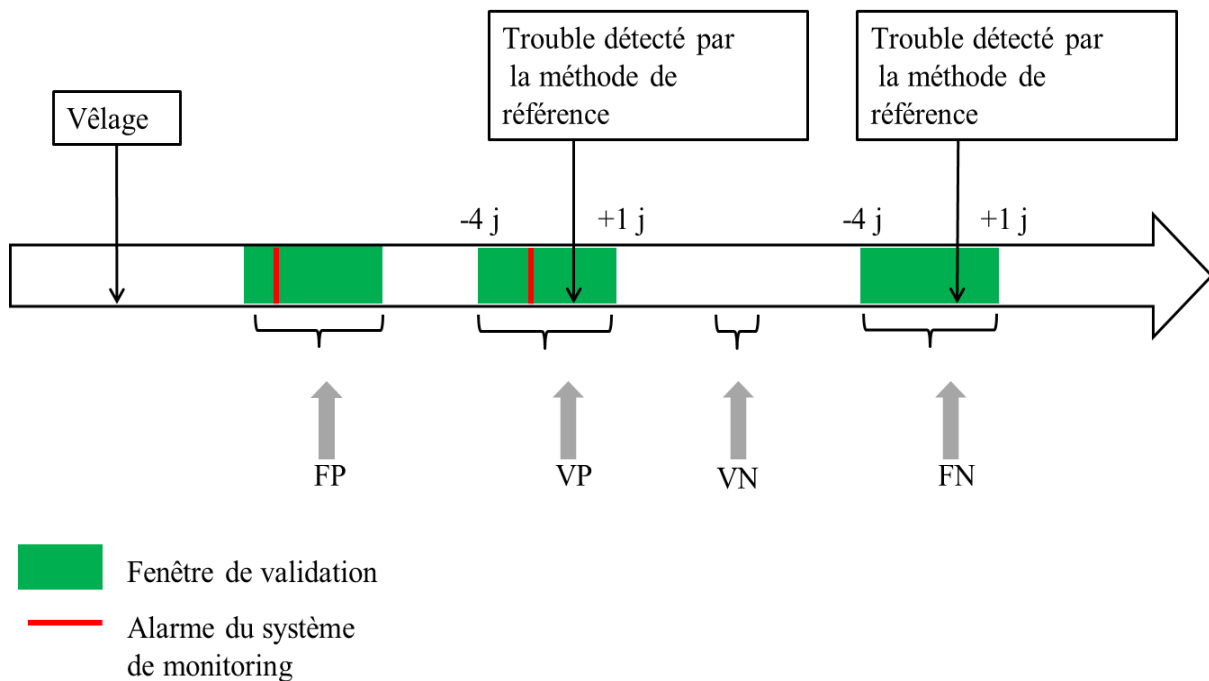


Figure 1-2 : principe de la comptabilisation des évènements pour le calcul de la valeur informative des alertes générées par un système de monitoring (FP=faux positif, VN=vrai négatif, VP=vrai positif, FN=faux négatif).

La définition des évènements vrais négatifs est parfois faite avec une fenêtre temporelle de taille différente de celle utilisée pour les autres évènements (vrais positifs ou faux négatifs), par exemple à l'échelle de la traite pour la détection des mammites. **Les fenêtres temporelles utilisées pour le calcul des valeurs informatives sont très variables entre études et parfois difficiles à définir.** Elles sont généralement choisies assez empiriquement en fonction des connaissances physiologiques sur la maladie d'intérêt (délai possible entre le début de l'infection et l'apparition des grumeaux dans le lait pour les mammites par exemple) et de l'applicabilité pour l'utilisateur des alertes (peu intéressant d'alerter l'éleveur sur une mammite après qu'il l'ait détecté via les grumeaux par exemple). Plusieurs tailles de fenêtres sont souvent testées dans une même étude (Adams et al., 2013).

La **précocité des alertes est également un aspect important.** En effet, un système produisant des alertes de santé n'est intéressant que s'il détecte des évènements avant ou *a*

minima en même temps que l'éleveur. La précocité des alertes n'est en général pas traitée ou alors présentée comme un résultat peu important dans les publications (Adams et al., 2013).

3.5.3.2 Statistiques opérationnelles pour l'évaluation des alertes

Lorsque l'on souhaite évaluer la valeur informative d'alertes de santé, **il est important de disposer d'indicateurs opérationnels permettant de juger de l'applicabilité des alertes pour l'éleveur**. En effet, un système d'alerte générant beaucoup d'alertes sur des animaux non malades, ou peu d'alertes sur des animaux malades, entraînera une perte de confiance de l'éleveur dans les alertes générées et une sous-utilisation du système d'alerte. La sensibilité et la spécificité ne reflètent que partiellement les performances d'un système d'alerte en conditions réelles d'utilisation car celles-ci dépendent également des conditions dans lesquelles il est placé, et notamment de la prévalence des troubles de santé dans l'élevage. Il existe donc des indicateurs comme la proportion d'animaux en alerte réellement malades (valeur prédictive positive) ou la proportion d'animaux non en alerte réellement sains (valeur prédictive négative). Des indicateurs comme l'Average Run Length (ARL) (Huybrechts et al., 2014) peuvent également être utilisés pour quantifier la durée moyenne entre 2 alertes en l'absence de trouble (ARL_0) et la durée moyenne entre la détection d'un trouble et la première alerte par le système (ARL_1). Le meilleur système d'alerte est celui qui permet un ARL_0 élevé et un ARL_1 court. Le taux de faux positifs générés par le système (Kamphuis et al., 2008, Miekley et al., 2013) est également un bon indicateur de son applicabilité en élevage. **Cependant, le calcul de ces indicateurs opérationnels n'est pas toujours mené dans les études s'intéressant à des systèmes de détection des troubles de santé, peut être en raison des difficultés méthodologiques rencontrées** (définition des événements vrais négatifs par exemple).

3.5.3.3 Choix d'une méthode de référence pertinente pour la détection des troubles de santé

L'évaluation de la valeur informative de systèmes d'alerte se fait généralement par classification croisée des événements de santé suivant une méthode de référence permettant

de déterminer l'état de santé « vrai » des animaux (sains ou malades) ainsi que celui « mesuré » par le système d'alerte à évaluer.

Tableau 1-4: statistiques relatives à la valeur informative d'un système d'alerte.

		Méthode de référence		Valeur informative ¹
		Malade	Non malade	
Alerte	oui	a	b	Se=a/(a+c)
	Non	c	d	Sp=d/(b+d)
				VPP=a/(a+b)
				VPN=d/(c+d)

¹Se= sensibilité, Sp=spécificité, VPP=valeur prédictive positive, VPN=valeur prédictive négative

Les difficultés pour évaluer la valeur informative d'un outil de monitoring de la santé comme Heatime-HR® sont de plusieurs ordres.

Tout d'abord, en raison de **la largeur du spectre de troubles potentiellement détectables par l'outil** : nous avons vu dans la partie 1.3.4 que peu d'études avaient tenté d'évaluer les potentialités de l'outil pour la détection des troubles de santé. Or les grandeurs mesurées par les capteurs portés par les animaux sont relativement non spécifiques et peuvent être impactées par une grande variété de troubles de santé (Radostits et al., 2007). Par conséquent, l'évaluation exhaustive de l'outil obligerait à multiplier de manière répétée dans le temps des tests correspondant aux méthodes de référence permettant de diagnostiquer chaque maladie potentiellement détectable (bactériologie pour les mammites, score de locomotion pour les boiteries, dosages biochimiques pour les troubles métaboliques), ce qui est coûteux. De plus, les troubles de santé exprimés cliniquement étant relativement rares en élevage laitier (Fourichon et al., 2001), les études doivent porter sur un nombre important d'animaux, notamment pour déterminer la sensibilité des détections avec précision. Deux stratégies d'évaluation de l'outil sont alors envisageables. Soit évaluer finement les performances de l'outil pour la détection de quelques maladies cibles, choisies par exemple sur la base de leur

fréquence ou de leur impact économique. Soit utiliser une méthode de référence qui soit plus imparfaite mais qui permette d'explorer d'emblée une large gamme de troubles de santé.

On doit également tenir compte de **l'imperfection des méthodes de référence** : en effet, la première méthode de référence qui vient à l'esprit pour évaluer un outil de monitoring des troubles de santé en élevage est de **l'évaluer par rapport à la méthode « classique » qui est la détection par observation visuelle par les éleveurs**. C'est une méthode utilisée couramment, notamment pour la détection des mammites (Rutten et al., 2013) ou des troubles hyperthermisants (Adams et al., 2013). Cependant des travaux ont montré que pour certaines maladies comme les boiteries (Fabian et al., 2014, Leach et al., 2010) ou les troubles respiratoires par exemple, **une faible partie des troubles mesurables** (25% des hyperthermies liées aux troubles respiratoires chez des taurillons en début d'engraissement) **était effectivement détectée par les éleveurs** (Timsit et al., 2011b). De plus, la qualité de la détection par observation visuelle est vraisemblablement variable entre éleveurs suivant leur expérience et les maladies auxquelles ils accordent une priorité, mais aussi variable dans le temps pour un éleveur donné, en fonction du temps dont il dispose, de l'attention qu'il peut accorder à ses animaux, de leurs conditions de logement (au pâturage ou en stabulation suivant le moment de l'année par exemple). On peut donc supposer que certains troubles de santé ne sont pas détectés par les éleveurs alors même qu'ils peuvent déclencher des alertes par le système de monitoring. **Ceci risque de conduire à une sous-estimation de la spécificité des alertes de santé si la détection par observation visuelle par l'éleveur est utilisée comme méthode de référence. Une méthode alternative devra donc être proposée pour évaluer la spécificité des alertes de santé. Cette méthode pourrait être un examen clinique vétérinaire approfondi d'animaux visant à déterminer leur statut sanitaire (sain ou malade) en parallèle de la détermination de leur statut vis-à-vis des alertes de santé (en alerte ou non en alerte).**

Enfin, **la méthode de référence peut être choisie pour ne détecter que certains types d'évènements**. Lorsqu'aucun Gold Standard n'est disponible ou que plusieurs méthodes de référence sont utilisées dans les études s'intéressant à une maladie (par exemple pour les mammites, où la référence peut être l'observation visuelle du lait, la concentration en cellules somatiques ou la bactériologie du lait), le choix de la méthode de référence utilisée peut être fonction de l'utilisation finale de l'outil à évaluer. **La question se pose de savoir ce que l'on veut détecter avec l'outil de monitoring évalué. En effet, ces outils de monitoring sont une aide, et les informations fournies doivent permettre de déboucher sur une prise de**

décision par l'éleveur. Ainsi, utiliser la concentration en cellules somatiques comme méthode de référence pour un outil susceptible de détecter les mammites peut conduire à un réglage très sensible des alertes mammites. L'éleveur devra alors savoir quelle conduite adopter en cas d'alerte, un traitement systématique des mammites subcliniques étant contraire aux objectifs de réduction des intrants médicamenteux en élevage. De plus, trop d'alertes risquent d'être générées, aboutissant à un surplus d'information à gérer pour l'éleveur et non à un allègement du travail physique et de la charge mentale. **Compte tenu du temps nécessaire à l'examen des animaux en alerte (dont une partie seulement est généralement examinée) et de la désorganisation provoquée par ces examens, il peut être pertinent de mettre au point des systèmes alertant l'éleveur uniquement pour les troubles les plus graves.** La gravité des troubles peut être définie suivant différents critères : cliniques (gravité et intensité des signes cliniques observés, éventuellement mort de l'animal lors de trouble aigu), zootechniques (impact du trouble sur la production laitière, réforme anticipée). **Les chutes de production laitière en particulier, mesurables facilement à l'aide de compteurs à lait ou sur les robots de traite, pourraient être un bon marqueur de l'existence et de la sévérité d'un trouble de santé (Bareille et al., 2003, Fourichon et al., 1999).** Les troubles les plus graves et devant signalés à l'éleveur en priorité sont ceux qui impactent la production laitière des animaux, puisqu'ils ont un impact économique direct par les pertes en lait engendrées et reflètent également un niveau de sévérité élevé du trouble de santé.

4. Objectif général, stratégie et questions scientifiques traitées, plan de thèse

L'objectif général de cette thèse était de produire des connaissances permettant de documenter l'intérêt de l'activité générale et des durées de rumination des vaches laitières, mesurées par l'appareil de monitoring Heatime-HR®, ou d'alertes générées à partir de ces données. Les résultats attendus étaient d'explorer les potentialités de l'appareil pour le phénotypage de caractères d'intérêt et d'évaluer la valeur informative d'alertes générées à partir de l'activité générale et des durées de rumination pour la détection des troubles de santé chez les vaches laitières. Les connaissances relatives à l'intérêt de l'activité et des durées de rumination pour la détection d'évènements de reproduction ont été jugées conséquentes au

cours de notre étude de la bibliographie et n'ont par conséquent pas fait l'objet de travaux dans cette thèse.

4.1 Intérêt de la mesure de l'activité générale et des durées de rumination pour le phénotypage

Nous avons vu dans le chapitre 1-3.3 que les mesures issues du système Heatime-HR® pouvaient potentiellement présenter un intérêt pour le phénotypage de l'intensité et de la cyclicité des chaleurs, des quantités ingérées ou de l'efficacité alimentaire des vaches, ainsi que pour le phénotypage des troubles de santé. Compte tenu des enjeux importants liés au phénotypage des quantités ingérées et de l'efficacité alimentaire, seul cet aspect a été abordé dans cette thèse. Le premier objectif de cette thèse était donc d'évaluer l'intérêt des durées de rumination fournies par Heatime-HR® pour estimer les quantités d'aliment ingérées par des vaches en lactation. La stratégie scientifique retenue pour répondre à cet objectif a été de comparer les qualités de prédiction des quantités ingérées d'un modèle de référence (modèle National Research Council 2001) et de ce même modèle de référence dans lequel les durées de rumination étaient ajoutées comme variables indépendantes. Les résultats obtenus sont présentés dans le **chapitre 2**.

4.2 Valeur informative d'alertes générées à partir de l'activité générale et des durées de rumination pour la détection des troubles de santé

Notre deuxième objectif était d'évaluer la valeur informative des données élémentaires de rumination et d'activité fournies par l'appareil Heatime-HR® pour la détection des troubles de santé. La stratégie scientifique retenue pour répondre à cet objectif a été de développer des alertes de santé, basées sur la détection de baisses d'activité générale ou des durées de rumination à l'aide d'algorithmes de types *cumulative sum*, puis à évaluer les performances des algorithmes développés pour la détection des principaux troubles cliniques rencontrés en élevage de vaches laitières. Les choix scientifiques principaux que nous avons faits étaient d'évaluer les alertes vis-à-vis de la détection d'une large variété de troubles de santé (cf chapitre 1 – 3.5.1), en utilisant comme indicateurs de performances des algorithmes la sensibilité, la spécificité, et les valeurs prédictives positives et négatives (chapitre 1 – 3.3.3.2),

calculées sur des fenêtres temporelles (chapitre 1 - 3.5.3.1) de taille identique (5 jours) pour tous les troubles de santé. La principale originalité de cette partie de la thèse a été d'évaluer les algorithmes développés par rapport à 3 méthodes de référence complémentaires (**chapitre 3**) qui présentaient a priori une qualité croissante, c'est-à-dire pour lesquelles le risque de mauvais classement des animaux (notamment classement des alertes en faux positifs au lieu de vrais positifs) était décroissant. L'objectif était notamment d'évaluer avec le plus de précision possible la spécificité des alertes de santé développées (chapitre 1 – 3.5.3.3). Les algorithmes de détection des troubles de santé ont donc été évalués successivement par rapport à la détection par observation visuelle de l'éleveur (méthode de référence numéro 1 : **chapitre 3.I**), par rapport à la détection par observation visuelle de l'éleveur avec prise en compte des baisses importantes de production laitière (méthode de référence numéro 2 : **chapitre 3.II**), en considérant que les baisses importantes de production laitière en dehors de toute détection de trouble par l'éleveur pouvaient résulter d'un trouble de santé non détecté par l'éleveur, et enfin les algorithmes de détection ont été évalués en utilisant comme méthode de référence un examen clinique vétérinaire systématique de vaches ayant présenté ou n'ayant pas présenté des chutes de rumination ou d'activité (méthode de référence numéro 3 : **chapitre 3.III**). Une autre originalité de notre travail a été d'évaluer la précocité des alertes de santé par rapport à la méthode de référence (chapitre 1 - 3.5.3.1) dans les chapitres 3.I et 3.II

Enfin, le **chapitre 4** consiste en une discussion générale du projet de thèse. Les principaux résultats de la thèse y sont rappelés et mis en perspective dans le cadre plus général de l'élevage de précision.

Références Bibliographiques

Adams, A. E., F. J. Olea-Popelka, and I. N. Roman-Muniz. 2013. Using temperature-sensing reticular boluses to aid in the detection of production diseases in dairy cows. *J Dairy Sci* 96:1549-1555.

Adin G, Solomon R, Nikbachat M, Zenou A, Yosef E, Brosh A, Shabtay A, Mabjeesh SJ, Halachmi I, Miron J. 2009. Effect of feeding cows in early lactation with diets differing in roughage-neutral detergent fiber content on intake behavior, rumination, and milk production. *J Dairy Sci* 92:3364-73.

Agreste. 2000. La production bovine française emploie l'équivalent de 30000 personnes à temps plein. <http://agreste.agriculture.gouv.fr/publications/cahiers/article/la-production-bovine-francaise>. Consultée le 16/06/2013.

Agreste. 2012. Productions animales. <http://agreste.agriculture.gouv.fr/IMG/pdf/Gaf12p145-163.pdf>. Consultée le 16/06/2013.

At-Taras, E. E. and S. L. Spahr. 2001. Detection and characterization of estrus in dairy cattle with an electronic heatmount detector and an electronic activity tag. *J Dairy Sci* 84:792-798.

Aungier, S. P., J. F. Roche, M. Sheehy, and M. A. Crowe. 2012. Effects of management and health on the use of activity monitoring for estrus detection in dairy cows. *J Dairy Sci* 95:2452-2466.

Bar, D. and R. Solomon. 2010. Rumination collars: What can they tell us. <http://www.precisiondairy2010.com/proceedings/s11bar.pdf>. Consultée le 10/04/2014.

Bareille, N., F. Beaudeau, S. Billon, A. Robert, and P. Faverdin. 2003. Effects of health disorders on feed intake and milk production in dairy cows. *Livest Prod Sci* 83:53-62.

Beauchemin, K. A., S. Zelin, D. Genner, and J. G. Buchanan-Smith. 1989. An automatic system for quantification of eating and ruminating activities of dairy cattle housed in stalls. *J Dairy Sci* 72:2746-2759.

Bewley, J. M. 2010. Precision dairy farming: Advanced analysis solutions for future profitability. <http://www.precisiondairy2010.com/proceedings/s1bewley.pdf>. Consultée le 20/07/2014.

Boichard, D. and M. Brochard. 2012. New phenotypes for new breeding goals in dairy cattle. *Animal* 6:544-550.

Burfeind, O., K. Schirmann, M. A. von Keyserlingk, D. M. Veira, D. M. Weary, and W. Heuwieser. 2011. Evaluation of a system for monitoring rumination in heifers and calves. *J Dairy Sci* 94:426-430.

Byskov, M. V., A. K. Schulze, M. R. Weisbjerg, B. Markussen, and P. Norgaard. 2014. Recording rumination time by a rumination monitoring system in jersey heifers fed grass/clover silage and hay at three feeding levels. *J Anim Sci* 92:1110-1118.

Chanvallon, A., S. Coyral-Castel, J. Gatien, J. M. Lamy, D. Ribaud, C. Allain, P. Clément, and P. Salvetti. 2014. Comparison of three devices for the automated detection of estrus in dairy cows. *Theriogenology*. <http://dx.doi.org/10.1016/j.theriogenology.2014.06.010>.

Cournut, S. and S. Chauvat. 2012. L'organisation du travail en exploitation d'élevage : Analyse de 630 bilans travail réalisés dans huit filières animales. *INRA Prod. Anim.* 25:101-112.

de Koning, C. J. A. M. 2010. Automatic milking-common practice on dairy farms. <http://www.precisiondairy2010.com/proceedings/s3dekonig.pdf>. Consultée le 20/07/2014.

de Mol, R. M., A. Keen, G. H. Kroeze, and J. M. F. H. Achten. 1999. Description of a detection model for oestrus and diseases in dairy cattle based on time series analysis combined with a kalman filter. *Comput Electron Agr* 22:171-185.

DeVries, T. J., K. A. Beauchemin, F. Dohme, and K. S. Schwartzkopf-Genswein. 2009. Repeated ruminal acidosis challenges in lactating dairy cows at high and low risk for developing acidosis: Feeding, ruminating, and lying behavior. *J Dairy Sci* 92:5067-5078.

Eastwood, C., D. Chapman, and M. Paine. 2004. Precision dairy farming-taking the microscope to dairy farm management.

EFSA (European Food Safety Authority). 2009. Scientific opinion of the panel on animal health and welfare on a request from european commission on welfare of dairy cows. Effects of farming systems on dairy cow welfare and disease. in *The EFSA Journal*. 1143. 1-38.

Chapitre 1 – Introduction générale

Elischer, M. F., M. E. Arceo, E. L. Karcher, and J. M. Siegford. 2013. Validating the accuracy of activity and rumination monitor data from dairy cows housed in a pasture-based automatic milking system. *J Dairy Sci* 96:6412-6422.

Fabian, J., R. A. Laven, and H. R. Why. 2014. The prevalence of lameness on new zealand dairy farms: A comparison of farmer estimate and locomotion scoring. *Vet J* 201:31-38.

Fitzpatrick, C. E., N. Chapinal, C. S. Petersson-Wolfe, T. J. DeVries, D. F. Kelton, T. F. Duffield, and K. E. Leslie. 2013. The effect of meloxicam on pain sensitivity, rumination time, and clinical signs in dairy cows with endotoxin-induced clinical mastitis. *J Dairy Sci* 96:2847-2856.

Fogsgaard, K. K., C. M. Rontved, P. Sorensen, and M. S. Herskin. 2012. Sickness behavior in dairy cows during escherichia coli mastitis. *J Dairy Sci* 95:630-638.

Fourichon, C., H. Seegers, N. Bareille, and F. Beaudeau. 1999. Effects of disease on milk production in the dairy cow: A review. *Prev Vet Med* 41:1-35.

Fourichon, C., F. Beaudeau, N. Bareille, H. Seegers. Incidence of health disorders in dairy farming systems in western France. 2001. *Liv Prod Sci* 68:157-170.

Fourichon, C., H. Seegers, N. Bareille, and F. Beaudeau. 2002. Impact économique des troubles de santé sous différentes logiques d'intensification de la production laitière en pays de la Loire. in *Rencontres Recherches Ruminants*. Paris, France.

Fricke, P. M., J. O. Giordano, A. Valenza, G. Lopes, Jr., M. C. Amundson, and P. D. Carvalho. 2014. Reproductive performance of lactating dairy cows managed for first service using timed artificial insemination with or without detection of estrus using an activity-monitoring system. *J Dairy Sci*. 97:2771-81.

Goldhawk, C., K. Schwartzkopf-Genswein, and K. A. Beauchemin. 2013. Technical note: Validation of rumination collars for beef cattle. *J Anim Sci* 91:2858-2862.

Gonzalez-Recio, O., J. E. Pryce, M. Haile-Mariam, and B. J. Hayes. 2014. Incorporating heifer feed efficiency in the Australian selection index using genomic selection. *J Dairy Sci* 97:3883-3893.

Hocquette, J.-F. and V. David. 2011. Les enjeux du phénotypage. *Bulletin Technique de l'Insemination Artificielle* 142.

Hogeveen, H., C. Kamphuis, W. Steeneveld, and H. Mollenhorst. 2010. Sensors and clinical mastitis--the quest for the perfect alert. *Sensors (Basel)* 10:7991-8009.

Holman, A., J. Thompson, J. E. Routly, J. Cameron, D. N. Jones, D. Grove-White, R. F. Smith, and H. Dobson. 2011. Comparison of oestrus detection methods in dairy cattle. *Vet Rec* 169:47-52.

Hostiou, N., C. Allain, S. Chauvat, A. Turlot, C. Pineau, and J. Fagon. 2014. L'élevage de précision : Quelles conséquences pour le travail des éleveurs. *INRA Prod. Anim.* 27:111-120.

Huybrechts, T., K. Mertens, J. De Baerdemaeker, B. De Ketelaere, and W. Saeys. 2014. Early warnings from automatic milk yield monitoring with online synergistic control. *J Dairy Sci* 97:3371-3381.

Idele (Institut de l'élevage). 2010. Référentiel travail en élevage bovin lait. <http://idele.fr/recherche/publication/idelesolr/recommends/referentiel-travail-en-elevages-bovins-lait.html>. Consultée le 16/06/2014.

Idele (Institut de l'élevage). 2012. Panorama des capteurs en élevage bovin. <http://idele.fr/recherche/publication/idelesolr/recommends/panorama-des-capteurs-en-elevage-bovin.html>. Consultée le 20/07//2014.

Isobe, N., T. Yoshimura, C. Yoshida, and T. Nakao. 2004. Incidence of silent ovulation in dairy cows during post partum period. *Dtsch Tierarztl Wochenschr* 111:35-38.

Kamphuis, C., R. Sherlock, J. Jago, G. Mein, and H. Hogeveen. 2008. Automatic detection of clinical mastitis is improved by in-line monitoring of somatic cell count. *J Dairy Sci* 91:4560-4570.

Kamphuis, C., B. DelaRue, C. R. Burke, and J. Jago. 2012. Field evaluation of 2 collar-mounted activity meters for detecting cows in estrus on a large pasture-grazed dairy farm. *J Dairy Sci* 95:3045-3056.

Kamphuis, C., B. Dela Rue, G. Mein, and J. Jago. 2013a. Development of protocols to evaluate in-line mastitis-detection systems. *J Dairy Sci* 96:4047-4058.

Kamphuis, C., E. Frank, J. K. Burke, G. A. Verkerk, and J. G. Jago. 2013b. Applying additive logistic regression to data derived from sensors monitoring behavioral and physiological characteristics of dairy cows to detect lameness. *J Dairy Sci* 96:7043-7053.

Laborde, A. 2012. Tic et agriculture. Appropriation des dispositifs numériques et mutations des organisations agricoles. ed. L'Harmattan, Paris. 242 pages.

Leach, K. A., H. R. Whay, C. M. Maggs, Z. E. Barker, E. S. Paul, A. K. Bell, and D. C. Main. 2010. Working towards a reduction in cattle lameness: 1. Understanding barriers to lameness control on dairy farms. *Res. Vet. Sci.* 89:311-317.

Lopez, H., L. D. Satter, and M. C. Wiltbank. 2004. Relationship between level of milk production and estrous behavior of lactating dairy cows. *Anim. Reprod. Sci.* 81:209-223.

Lukas, J. M., J. K. Reneau, R. Wallace, D. Hawkins, and C. Munoz-Zanzi. 2009. A novel method of analyzing daily milk production and electrical conductivity to predict disease onset. *J Dairy Sci* 92:5964-5976.

Miekley, B., E. Stamer, I. Traulsen, and J. Krieter. 2013. Implementation of multivariate cumulative sum control charts in mastitis and lameness monitoring. *J Dairy Sci* 96:5723-5733.

Miguel-Pacheco, G. G., J. Kaler, J. Remnant, L. Cheyne, C. Abbott, A. P. French, T. P. Pridmore, and J. N. Huxley. 2014. Behavioural changes in dairy cows with lameness in an automatic milking system. *Appl Anim Behav Sci* 150:1-8.

Mollenhorst, H., P. P. J. van der Tol, and H. Hogeveen. 2010. Somatic cell count assessment at the quarter or cow milking level. *J Dairy Sci* 93:3358-3364.

Neves, R. C., K. E. Leslie, J. S. Walton, and S. J. Leblanc. 2012. Reproductive performance with an automated activity monitoring system versus a synchronized breeding program. *J Dairy Sci* 95:5683-5693.

Palmer, M. A., G. Olmos, L. A. Boyle, and J. F. Mee. 2010. Estrus detection and estrus characteristics in housed and pastured holstein-friesian cows. *Theriogenology* 74:255-264.

Pryce, J. E., W. J. Wales, Y. de Haas, R. F. Veerkamp, and B. J. Hayes. 2014. Genomic selection for feed efficiency in dairy cattle. *Animal* 8:1-10.

Radostits, O. M., C. C. Gay, K. W. Hinchcliff, and P. D. Constable. 2007. *Veterinary Medicine: A textbook of the diseases of cattle, horses, sheep, pigs, and goats.* 10th ed. Saunders Elsevier, New York.

Ranasinghe, R. M. S. B. K., T. Nakao, K. Yamada, and K. Koike. 2010. Silent ovulation, based on walking activity and milk progesterone concentrations, in Holstein cows housed in a free-stall barn. *Theriogenology* 73:942-949.

Roelofs, J. B., F. J. van Eerdenburg, N. M. Soede, and B. Kemp. 2005. Pedometer readings for estrous detection and as predictor for time of ovulation in dairy cattle. *Theriogenology* 64:1690-1703.

Rutten, C. J., A. G. Velthuis, W. Steeneveld, and H. Hogeveen. 2013. Invited review: Sensors to support health management on dairy farms. *J Dairy Sci* 96:1928-1952.

Schirmann, K., M. A. von Keyserlingk, D. M. Weary, D. M. Veira, and W. Heuwieser. 2009. Technical note: Validation of a system for monitoring rumination in dairy cows. *J Dairy Sci* 92:6052-6055.

Schirmann, K., N. Chapinal, D. M. Weary, W. Heuwieser, and M. A. von Keyserlingk. 2012. Rumination and its relationship to feeding and lying behavior in holstein dairy cows. *J Dairy Sci* 95:3212-3217.

Shipka, M. P. 2000. A note on silent ovulation identified by using radiotelemetry for estrous detection. *Appl Anim Behav Sci* 66:153-159.

Shrestha, H. K., T. Nakao, T. Higaki, T. Suzuki, and M. Akita. 2004. Resumption of postpartum ovarian cyclicity in high-producing holstein cows. *Theriogenology* 61:637-649.

Soriani, N., E. Trevisi, and L. Calamari. 2012. Relationships between rumination time, metabolic conditions, and health status in dairy cows during the transition period. *J Anim Sci* 90:4544-4554.

Soriani, N., D. Bar, L. Calamari, and G. Tadini. 2013. Rumination time: An indicator of health status and welfare condition in Proc. Precision Dairy Conference, Rochester, Minnesota. Pages 91-92

Steensels, M., C. Bahr, D. Berckmans, A. Antler, E. Maltz, and I. Halachmi. 2012. Detection of early lactation ketosis by rumination and other sensors. http://www.eaap.org/Previous_Annual_Meetings/2012Bratislava/Papers/Published/27_Steensels.pdf. Consultée le 20/07/2014.

Sun, Z., S. Samarasinghe, and J. Jago. 2010. Detection of mastitis and its stage of progression by automatic milking systems using artificial neural networks. *J Dairy Res* 77:168-175.

Timsit, E., S. Assie, R. Quiniou, H. Seegers, and N. Bareille. 2011a. Early detection of bovine respiratory disease in young bulls using reticulo-rumen temperature boluses. *Vet J* 190:136-142.

Timsit, E., N. Bareille, H. Seegers, A. Lehebel, and S. Assie. 2011b. Visually undetected fever episodes in newly received beef bulls at a fattening operation: Occurrence, duration, and impact on performance. *J Anim Sci* 89:4272-4280.

Van Hertem, T., E. Maltz, A. Antler, C. E. Romanini, S. Viazzi, C. Bahr, A. Schlageter-Tello, C. Lokhorst, D. Berckmans, and I. Halachmi. 2013. Lameness detection based on multivariate continuous sensing of milk yield, rumination, and neck activity. *J Dairy Sci* 96:4286-4298.

Villaume, S. and N. Delame. 2009. Essor des sociétés agricoles : Un recours accru au salariat et aux prestataires de services. INSEE premiere 1243.

Wathes, C. M., H. H. Kristensen, J. M. Aerts, and D. Berckmans. 2008. Is precision livestock farming an engineer's daydream or nightmare, an animal's friend or foe, and a farmer's panacea or pitfall? *Comput Electron Agr* 64:2-10.

Annexe du chapitre 1

Annexe 1-1: description détaillée du fonctionnement de l’outil Heatime-HR®.

L’outil support de cette thèse, Heatime-HR®, mesure en continu **l’activité générale et la rumination** des animaux équipés et fournit des alertes de chaleurs et de santé à partir de ces données. Heatime-HR® a été développé par une société Israélienne (SCR Engineers Ltd, Netanya, Israel), est assemblé et commercialisé en Europe par une entreprise Italienne (Milkline, Piacenza, Italy) et distribué en France par une entreprise française de sélection animale (Evolution, Rennes, France) dans le cadre de son activité de fourniture de solutions de monitoring. Dans le système Heatime-HR®, les vaches sont équipées de colliers qui comportent un boîtier en plastique situé en partie distale de l’encolure et en contact avec les masses musculaires du 1/3 supérieur de l’encolure. Le boîtier contient 2 capteurs : 1 activimètre qui mesure les mouvements de l’animal dans les 3 dimensions de l’espace et 1 microphone qui mesure les sons. Les données brutes issues des capteurs sont ensuite traitées automatiquement (filtrages de certaines fréquences pour les sons par exemple) pour donner une donnée d’activité (sans unité) et une durée de rumination (minutes) avec une résolution de 2 minutes. Ces données sont ensuite agrégées et stockées pour fournir une valeur d’activité et une durée de rumination toutes les 2 heures (12 valeurs par jour pour chaque donnée), qui sont les données accessibles aux utilisateurs. Les capteurs ne sont pas synchronisés entre eux et, dans une ferme donnée, une même période de 2 heures affichée par l’appareil pour 2 animaux différents peut ne pas correspondre exactement à la même période de temps. Les données sont ensuite transférées périodiquement du boîtier à un boîtier électronique situé par exemple dans la laiterie. Ce boîtier stock et permet de visualiser les données de chaque vache équipée sous forme de courbes. Les données peuvent également être visualisées sur un ordinateur dans une autre configuration d’installation. L’outil fournissait initialement uniquement les données brutes et des alertes pour la détection des chaleurs. Depuis mi-2011, il applique sur les données des algorithmes pour la détection des chaleurs et des troubles de santé, et fournit à l’éleveur des listes de vaches détectées en chaleur ou en alerte de santé. Les algorithmes de détection sont la propriété de SCR Engineers Ltd et ne sont donc pas connus dans le détail mais ils sont basés sur le calcul de déviations dans les valeurs prises pour une vache donnée

dans les 10 dernières heures par rapport aux valeurs prises pour cette même vache dans les 10 derniers jours. Chaque vache est ainsi son propre témoin. Un algorithme détecte les hyperactivités liées aux chaleurs et les vaches peuvent apparaître sur la liste des vaches en alerte de santé pour 3 raisons ; si elles ont ruminé moins de 300 minutes dans les dernières 24 heures, si l'algorithme de sous-activité est inférieur à -30, si l'algorithme de sous-rumination est inférieur à -30, ou si l'algorithme Index de Santé est supérieur à 86. Ces seuils sont ceux recommandés par le fabricant mais sont paramétrables par l'éleveur. La technique de transfert des données a évolué depuis le début des tests Heatime-HR® en France (2009). Initialement, le transfert des données était effectué par infrarouge lors du passage des vaches sous des antennes situées en entrée ou sortie de salle de traite ou au-dessus des abreuvoirs. Ce système garantissait le transfert des données et la génération des alertes au moins 1 fois par 24 heures, voir 2 fois dans la majorité des élevages. Depuis début 2013, les données sont transmises « en continu » par ondes wifi tout au long de la journée. Les alertes peuvent donc être générées plus de 2 fois par jour. Si les capteurs semblent avoir peu évolués depuis 2009, la partie logiciel de l'outil a connu plusieurs mises à jour sans que le détail des modifications effectuées (« habillage » des écrans uniquement ou modification des algorithmes ?) ne soit connu du distributeur. En plus des alertes de chaleur et de santé, Heatime-HR® joue également le rôle de planning de reproduction puisque l'éleveur, s'il rentre les dates de vêlage, d'IA et les résultats des diagnostics de gestation peut avoir accès à ces informations ainsi qu'à des informations supplémentaires (vaches en anoestrus par exemple) sous forme graphique ou sous forme de listes de vaches.

**Chapitre 2. Evaluation de l'intérêt de la mesure des durées
de rumination pour la prédiction des quantités
ingérées chez des vaches laitières en lactation**

1. Résumé du chapitre 2

L'objectif de cette étude était d'évaluer l'impact en termes de prédiction de l'ingestion, de l'inclusion des durées de rumination dans le modèle NRC (National Research Council) 2001 de prédiction des quantités de matières sèches ingérées chez les vaches laitières.

Quarante et une vaches Holstein recevant une ration complète identique ont été suivies pendant 10 semaines dans une ferme expérimentale de l'ouest de la France (Les Trinottières, Maine et Loire). Les durées de rumination individuelles étaient mesurées en continu et les quantités individuelles ingérées étaient mesurées quotidiennement. L'exactitude et la précision du modèle NRC 2001, qui utilise le poids vif, la quantité de lait standardisé à 4% de matière grasse et la semaine de lactation comme variables indépendantes, a été comparée à l'exactitude et à la précision du même modèle incluant la durée de rumination comme variable indépendante supplémentaire.

La durée de rumination avait un effet significatif sur les quantités de matière sèche ingérées dans le modèle développé, mais la valeur de son estimé était faible (0,031 kg/h). L'erreur moyenne de prédiction était très proche dans les 2 modèles (RMSPE = 1,70 kg/j pour le modèle sans rumination et 1,68 kg/j pour le modèle avec rumination), de même que les autres statistiques (R^2 , biais linéaire, erreur aléatoire et coefficient de corrélation-concordance) utilisées pour comparer les 2 modèles dans cette étude.

Ces résultats indiquent qu'il n'y avait pas de gain majeur d'exactitude ou de précision pour la prédiction des quantités ingérées lorsque la durée de rumination était incluse dans le modèle NRC 2001.

2. Interpretative summary

Added value of rumination time for the prediction of dry matter intake in lactating dairy cows

By Clément P. et al., p xxx. Predicting or measuring dry matter intake (DMI) in dairy cows is a major concern in diet formulation. This study aimed to investigate the change in DMI prediction provided by the inclusion of individual rumination time, a recently available on-farm measure, in the National Research Council (2001) DMI prediction model. The rumination time estimate was significant in the model developed. However, the prediction did not gain in precision or accuracy. Our results indicate that there is no benefit in measuring rumination time for improving the prediction of the DMI of individual dairy cows.

Short communication: added value of rumination time for the prediction of dry matter intake in lactating dairy cows

Accepté. J Dairy Sci. 2014. [http://dx.doi.org/ 10.3168/jds.2013-7860](http://dx.doi.org/10.3168/jds.2013-7860)

P. Clément*†, R. Guatteo*†, L. Delaby‡, B. Rouillé§, A. Chanvallon§, J. M. Philipot# and N. Bareille*†

*LUNAM Université, Oniris, UMR BioEpAR, CS 40706, 44307 Nantes, France

†INRA, UMR1300 BioEpAR, CS 40706, 44307 Nantes, France

‡ INRA, AgroCampus Ouest, UMR1348 Pegase, F-35590 Saint-Gilles, France

§Institut de l'Élevage, 149 rue de Bercy, 75595 Paris, France

#Evolution, 69 rue de la Motte Brûlon, 35700 Rennes, France

Corresponding author: pierre.clement@oniris-nantes.fr

3. Abstract

The objective of the current study was to quantify the change in the prediction of dry matter intake (DMI) resulting from the inclusion of rumination time (RT) in the 2001 National Research Council (NRC) DMI prediction model. Forty-one Holstein cows fed the same total mixed ration were involved in a 10-week study. Individual DMI were measured daily. The accuracy and precision of the original NRC prediction model, based on body weight, fat corrected milk and week of lactation as independent variables, was compared with the accuracy and precision of the same model with RT as an additional independent variable. The RT estimate was significant in the model developed but had a low value (0.031 kg/h). Root mean square prediction errors were very similar in the two models (1.70 and 1.68 kg/d) as were the other indicators (R^2 , linear bias, random error and concordance correlation coefficient) selected to compare the models in this study. These results indicate that there was no gain in DMI prediction precision or accuracy when RT was included in the NRC model.

Key words: dry matter intake, dairy cow, rumination, intake prediction

4. Short communication: added value of rumination time for the prediction of dry matter intake in lactating dairy cows

Prediction of DMI is important for the formulation of nutritionally and economically optimized diets in dairy cows. There have been extensive research efforts to develop a DMI prediction model that would be accurate over lactation for a large variety of diets. It has, however, proven difficult to create a general and accurate model due to the complex interactions between animal and feed factors (Huhtanen et al., 2011) which are poorly understood. Many models have been developed adapted to a particular country or a system for calculating nutritional requirements, but none of them can be used worldwide because they are insufficiently flexible to be accurate and precise in multiple different farming and feeding systems. Some models only include animal factors such as milk yield, stage of lactation or BW (NRC, 2001; Fox et al., 2004), while others include feed factors such as digestibility of organic matter, NDF content, and feed DM content (Lewis, 1981; Huhtanen et al., 2011), in

addition to animal factors. Simple models with few input variables lack accuracy, with mean biases of 1 or 2 kg DMI/day. On the other hand, more complex models may be more accurate and have mean biases lower than 1 kg DMI/day (Huhtanen et al., 2011; Krizsan et al., 2014) but are less suitable for on-farm application as they require large amounts of input data concerning diet characteristics.

In recent years, a rumination sensor has been developed (HR-Tag, SCR Engineers Ltd, Netanya, Israel). It enables the continuous recording of individual rumination times (**RT**) with high accuracy (Schirmann et al., 2009). RT is influenced by diet composition, for example, NDF content (Adin et al., 2009). Thus, RT could be seen as an on-animal measure that is linked to diet composition and might improve existing DMI prediction models. Many studies have evaluated the impact of diet composition and diet particle size on DMI or RT, but there are limited data concerning the relationship between these two variables for cows fed the same TMR. Hasegawa et al. (1997) suggested that RT may be used to assess DMI and Krause et al. (2002) found a positive relationship between long particle DMI and RT. More recently, Schirmann et al. (2012) found that daily RT and DMI were not correlated across cows but were weakly negatively correlated within cow, in 42 dry Holstein cows. Moreover, little attention has been paid to the relationship between DMI and RT during lactation. A positive relationship would be expected, as more RT may be necessary to process a growing amount of food.

The objective of this study was to investigate the putative gain in terms of DMI prediction accuracy and precision provided by the inclusion of RT in an existing and widely used prediction model developed in the USA by the National Research Council (NRC) (2001).

Forty-one (14 primiparous and 27 multiparous) Holstein cows at the Les Trinottières experimental farm (47°34'N; 0°23'W, Maine-et-Loire, France) were included in the study. The cows were housed in cubicles and fed a TMR for a 10-week period between April and July 2011. The TMR was formulated according to the recommendations provided by the French National Institute for Agricultural Research (INRA, 2010) and cows were fed ad libitum once a day at approximately 1300 h to achieve 10% refusals on an as-fed basis. The diet was 71.8% corn silage, 1.7% straw and 26.5% concentrate mix (rapeseed based) on a DM basis (CP = 14.1%, NDF = 43.2% of DM). Individual amounts of feed offered and refusals

were weighed daily. TMR samples were taken daily and dried at 105°C for 24 h to determine the DM content. Daily individual DMI was calculated for each cow by subtracting the refusals from the TMR offered.

To monitor rumination time, all cows were fitted with a rumination logger (HR-Tag, SCR Engineers Ltd, Netanya, Israel) supported by a neck collar. The logger contains a microphone that records RT continuously and summarizes it into 2 h intervals (Schirmann et al., 2009). Two-hour interval RT were summed between 1200 h on day n and 1200 h on day n+1 to obtain individual daily RT.

Cows were milked twice daily at 0700 and 1700 h and daily yield was recorded using milk meters. Samples were collected weekly at four consecutive milkings and assessed for fat content using infrared analyses (AOAC, 1997), results being summarized for each animal as a weekly mean. FCM was calculated daily as $FCM = (\text{kg of milk} \times 0.4) + (\text{kg of milk} \times \text{fat\%}_{\text{week}} / 100 \times 15)$ where $\text{fat\%}_{\text{week}}$ is the weekly mean milk fat percentage (Gaines, 1928). Daily FCM and RT values were averaged to obtain weekly means.

Individual BW were measured every 3 to 5 weeks (weeks -3, 2, 7 and 10 of the trial) at the same time of the day. Weekly BW were calculated by linear interpolation between two measurements.

The experimental unit was the cow week. In a first step, RT variance was decomposed using SAS proc VARCOMP (SAS version 9.2; SAS Institute Inc., Cary, NC) in order to estimate the variability linked to animal number, week and residual variability, for descriptive purposes. In a second step, two models were compared, NRC and NRC_{RT} :

$$DMI = (0.372 \times FCM + 0.0968 \times BW^{0.75}) \times (1 - \exp\{-0.192 \times (WL + 3.67)\}) \quad (\text{NRC})$$

$$DMI = (0.372 \times FCM + 0.0968 \times BW^{0.75}) \times (1 - \exp\{-0.192 \times (WL + 3.67)\}) + a \times RT \quad (\text{NRC}_{RT})$$

where DMI is the average daily DMI for one week, **WL** is the week of lactation, FCM is expressed in kg/d, BW in kg, and RT in h/d. The constant “a” was estimated using the proc REG (SAS version 9.2; SAS Institute Inc., Cary, NC).

Two criteria were calculated to assess and compare the goodness of fit of the models: the R^2 value of the regression between observed (O_i) and predicted (P_i) values; and the root mean square prediction error (RMSPE), calculated as $\text{RMSPE} = \text{square root}(\sum(O_i - P_i)^2/n)$. The statistical significance of the mean bias and linear bias were estimated as proposed by St-Pierre (2003). The mean square prediction error was divided into three components, namely, the mean bias, linear bias (the deviation of the slope from 1), and random variation around the regression line (Bibby and Toutenberg, 1977). Further, the synthetic index for assessing the accuracy and precision, the concordance correlation coefficient (CCC), was calculated for each model as proposed by Lin (1989).

Due to technical problems with some of the rumination monitoring devices, 13 animals only had RT recordings for 4 to 9 weeks of the trial. Weeks without RT data were excluded from the analyses. Descriptive statistics of the 371 cow weeks included in the analyses are provided in table 2-1. There was no linear relationship between the mean time spent ruminating and the mean DMI calculated at the cow level over the trial (figure 2-1). The mean RT measured in this study was 7.5 ± 0.1 h/d. For individual cows, the mean RT over the trial ranged from 3.8 to 9.7 h/d which is wider than that reported by Soriani et al. (2012) in early lactation (from 6.7 to 11.4 h/d). In our study, we observed a high variability in RT between cows. Indeed, across cows, 84.1% of RT variability was linked to animal number and 0.5% to week, 15.4% being residual variability. The variability linked to animal and to week were respectively 72.4% and 12.3% for DMI, and 85.7% and 3.3% for FCM. The CV for RT (22.7%) was higher than the CV obtained for DMI or FCM (10.4 and 14.1% respectively). At the cow level, the between-week CV was on average 8.9% (range = 3.6 to 19.2%) for RT, 5.2% (range = 1.9 to 10.4%) for DMI and 5.0% (range = 1.8 to 10.4%) for FCM. This result shows that for a given cow, RT was less repeatable between weeks than DMI or FCM. This is surprising as the main factors known to influence rumination activity, such as diet digestibility, diet fibrosity index (Sauvant et al., 1990), NDF intake (Adin et al., 2009), forage quality (Welch and Smith, 1970), heat stress (Soriani et al., 2013) and regrouping (Schirmann et al., 2011), were stable in our study. This higher CV might partly be due to imprecision in the RT measurements. Even though good correlations ($r = 0.93$, $R^2 = 0.87$) were obtained between HR-Tag measures and visual observation (Schirmann et al., 2009), in field conditions the RT measure is probably less accurate and precise than approved milk yield recording devices. Within-cow variability of RT might be a problem for the practical implementation of a model with RT as an input.

Indeed, past or present RT would be used in the model as future RT would not be known at the time of prediction, as occurs when milk yield is used as an input variable (Ingvarsen, 1994).

Table 2-1. Descriptive statistics (mean, standard deviation, minimum, maximum) of the variables included into the models (n = 371 cow weeks)

Criterion	Mean	SD	Min	Max
DMI, kg/d	25.9	2.66	19.4	33.1
FCM, kg/d	34.0	4.76	22.9	47.7
Rumination time, h/d	7.5	1.70	2.7	11.2
DIM, d	210	43.6	85	308
BW, kg	697	65.8	538	865

Many authors have studied the effect of differences in diet composition on RT but our study is the first to evaluate the relationship between DMI and RT for lactating cows fed the same TMR. As a basis, we used the NRC model (NRC, 2001), which has been widely used as a reference for the evaluation of new models (Halachmi et al., 2004; Shah and Murphy, 2006; Huhtanen et al., 2011). On the one hand, it is easy to implement because as input data it only uses “animal factors which would be easily measured or known” (NRC, 2001). On the other hand, the NRC model was shown to clearly over predict DMI when the diets used for evaluation were composed of silages other than those used for the development of the model (Huhtanen et al., 2011). In this case, inaccuracies of the model could be due to the fact that it does not include diet related factors. Another limitation of this model is that milk yield is used as an input variable. The milk yield is a function of the genetic merit of the cow and the quality of the diet. As a consequence, when current diet deviates from the economically optimal diet, using actual milk yield as an input variable can result in seriously biased intake predictions (Friggens et al., 1998), as discussed by Krizsan et al. (2014).

In the NRC_{RT} model, the estimate for RT was low (0.031 kg/h, $P < 0.01$) but significantly different from zero (table 2-2). This result is not consistent with the result obtained by Schirmann et al. (2012), who found no relationship between daily RT and DMI during the dry period. In NRC and NRC_{RT} models, there was a tendency for residuals to decrease when WL increased (-0.07 kg/WL, $P < 0.001$). Thus, in both models DMI was underestimated when WL was lower than 30 and overestimated otherwise. However, as this relationship was similar in the two models, it was still relevant to compare their goodness of fit. Residuals were not related to other independent variables.

Table 2-2. Statistical values obtained for NRC and NRC_{RT} models fitted to weekly DMI of individual cows, and accuracy and precision of the 2 models

Model	RT Estimate (SE)	Mean Bias (kg/d)	Linear Bias	RMSPE ¹ (kg/d)	R ²	CCC ₂	Error distribution ³		
							M	S	R
NRC (2001)	-	0.22 $P = 0.01$	-0.06 $P > 0.05$	1.70	0.60	0.76	0.02	0.00	0.98
NRC _{RT}	0.031 (0.0114)	-0.02 $P > 0.05$	-0.06 $P > 0.05$	1.68	0.60	0.76	0.00	0.01	1.00

¹RMSPE = root mean square prediction error expressed in kilograms per day.

²CCC = concordance correlation coefficient according to Lin (1989).

³Error distribution: M = mean, S = slope, R = random.

The accuracy and precision of the NRC and NRC_{RT} models are compared based on R², mean bias and RMSPE calculations summarized in table 2-2. The goodness of fit in our study (R² = 0.60 for NRC model) is better than that obtained by Shah and Murphy (2006) with a larger data set (R² = 0.42) in early lactation. The difference in mean days in milk (210 d vs. 35 d respectively) may partly explain the poorer fit of Shah and Murphy (2006), as during early lactation DMI increases less quickly than milk yield. Accordingly, it is more difficult to predict DMI accurately and precisely during this period. The number of data sets used (one data set in our study vs. two data sets originating from two farms in the study by Shah and Murphy (2006)) and the differences in diet composition may also contribute to the difference in goodness of fit. On the other hand, compared with our results, Huhtanen et al. (2011) obtained a higher R² (0.73). This may be attributable to the differences in diet composition between the two studies.

The plot used for NRC_{RT} bias decomposition is displayed in figure 2-2. The mean bias value was higher in the NRC model (0.22 kg/d, $P < 0.05$) than in the NRC_{RT} model (-0.02 kg/d, $P > 0.05$). This result was due to the fact that, in the NRC_{RT} model, the constant “a” was calculated to center the distribution of residuals around zero. Thus, a lower mean bias was to be expected for this model. RMSPE values were very similar (1.70 and 1.68 kg/d respectively) in the two models (table 2-2), as were R^2 values (0.60). The very high proportion of random error (0.98 for NRC model and 1.00 for NRC_{RT} model) indicates the absence of systematic errors in both models. Thus, the indicators selected for this study were consistent and confirmed that there was no gain in accuracy and precision in DMI prediction when RT was added to the NRC model.

In conclusion, in this study, the RT estimate (0.031 kg/h) was significant in the DMI prediction model developed, but the addition of RT to the model had no effect on the accuracy and precision indicators studied. Moreover, the variability of RT between weeks for a given cow fed the same TMR might make it difficult in practice to include RT in a DMI prediction model.

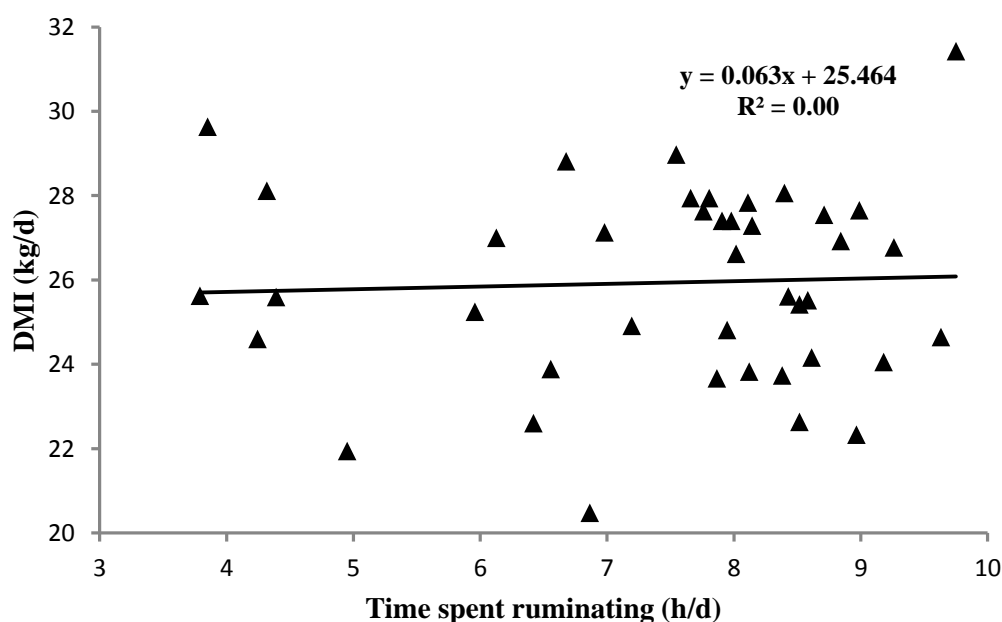


Figure 2-1. Scatter plot representing average time spent ruminating (h/d) against average DMI (kg/d) over the trial (n = 41 cows).

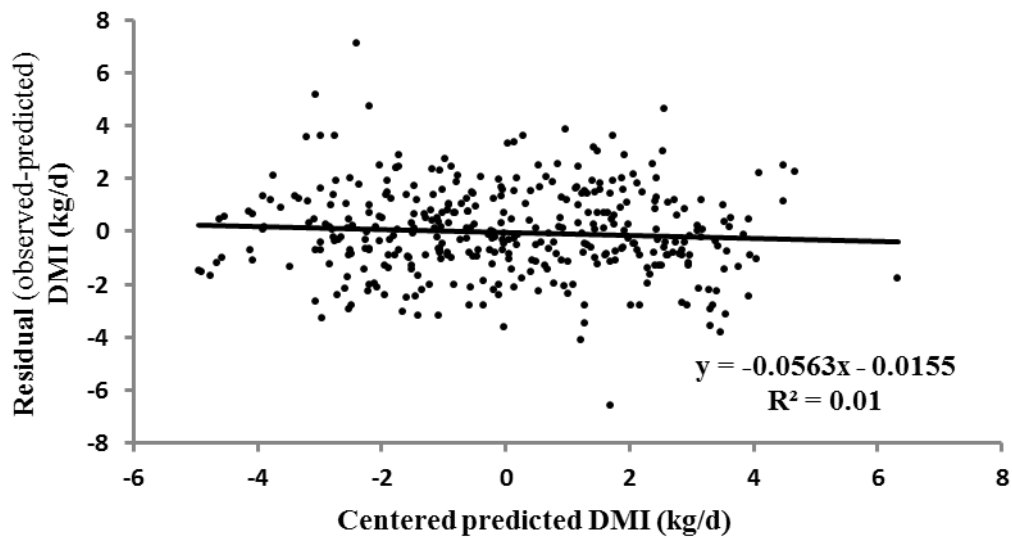


Figure 2-2. Scatter plot representing residual DMI (kg/d) against centered predicted DMI (kg/d) obtained with NRC_{RT} model (n = 371 cow weeks).

5. Acknowledgments

The authors thank Trinottières farm staff for their help, in particular with data recording. “Chambre d’Agriculture du Maine et Loire” and “Institut de l’élevage” are also acknowledged for their financial and technical support. Help from Anne Lehebel with statistical analyses was also very much appreciated.

References

- Adin, G., R. Solomon, M. Nikbachat, A. Zenou, E. Yosef, A. Brosh, A. Shabtay, S. J. Mabjeesh, I. Halachmi, and J. Miron. 2009. Effect of feeding cows in early lactation with diets differing in roughage-neutral detergent fiber content on intake behavior, rumination, and milk production. *J. Dairy Sci.* 92:3364-3373.
- AOAC (Association of Official Analytical Chemists). 1997. *Official Methods of Analysis*. 16th ed. AOC International, Gaithersburg, MD.
- Bibby, J. and H. Toutenberg 1977. *Prediction and improved estimation in linear models*. John Wiley & Sons Ltd., Chichester, UK.
- Fox, D. G., L. O. Tedeschi, T. P. Tytlutki, J. B. Russell, M. E. Van Amburgh, L. E. Chase, A. N. Pell, and T. R. Overton. 2004. The Cornell Net Carbohydrate and Protein System model for evaluating herd nutrition and nutrient excretion. *Anim Feed Sci Technol* 112:29-78.
- Friggens, N. C., G. C. Emmans, I. Kyriazakis, J. D. Oldham, and M. Lewis. 1998. Feed intake relative to stage of lactation for dairy cows consuming total mixed diets with a high or low ratio of concentrate to forage. *J. Dairy Sci.* 81:2228-2239.
- Gaines, W. L. 1928. An efficiency formula for dairy cows. *Science* 67:353-354.
- Halachmi, I., Y. Edan, U. Moallem, and E. Maltz. 2004. Predicting feed intake of the individual dairy cow. *J. Dairy Sci.* 87:2254-2267.
- Hasegawa, N., A. Nishiwaki, K. Sugawara, and I. Ito. 1997. The effects of social exchange between two groups of lactating primiparous heifers on milk production, dominance order, behavior and adrenocortical response. *Appl Anim Behav Sci* 51:15-27.
- Huhtanen, P., M. Rinne, P. Mantysaari, and J. Nousiainen. 2011. Integration of the effects of animal and dietary factors on total dry matter intake of dairy cows fed silage-based diets. *Animal* 5:691-702.
- Ingvartsen, K. L. 1994. Models of voluntary food intake in cattle. *Livest. Prod. Sci.* 39:19-38.

INRA (Institut National de la Recherche Agronomique). 2010. Alimentation des bovins ovins et caprins. Besoins des animaux-Valeur des aliments-Tables INRA 2007. 2010 ed. Quae, Versailles, France.

Krause, K. M., D. K. Combs, and K. A. Beauchemin. 2002. Effects of forage particle size and grain fermentability in midlactation cows. II. Ruminal pH and chewing activity. *J. Dairy Sci.* 85:1947-1957.

Krizsan, S. J., A. Sairanen, A. Hojer, and P. Huhtanen. 2014. Evaluation of different feed intake models for dairy cows. *J. Dairy Sci.* 97:2387-2397.

Lewis, M. 1981. Equations for predicting silage intake by beef and dairy cattle. Pages 35-36 in Summary of the 6th Silage Conference Edinburg, UK.

Lin, L. I. 1989. A concordance correlation coefficient to evaluate reproducibility. *Biometrics* 45:255-268.

NRC (National Research Council). 2001. Nutrient Requirements of Dairy Cattle. 7th revised edition ed. National Academy Press. Washington, DC, USA.

Sauvant, D., J. P. Dulphy, and B. Michalet-Doreau. 1990. The fibrosity index of ruminant feed and diet. *INRA Prod. Anim.* 3:309-318.

Schirmann, K., M. A. von Keyserlingk, D. M. Weary, D. M. Veira, and W. Heuwieser. 2009. Technical note: Validation of a system for monitoring rumination in dairy cows. *J. Dairy Sci.* 92:6052-6055.

Schirmann, K., N. Chapinal, D. M. Weary, W. Heuwieser, and M. A. von Keyserlingk. 2011. Short-term effects of regrouping on behavior of prepartum dairy cows. *J. Dairy Sci.* 94:2312-2319.

Schirmann, K., N. Chapinal, D. M. Weary, W. Heuwieser, and M. A. von Keyserlingk. 2012. Rumination and its relationship to feeding and lying behavior in Holstein dairy cows. *J. Dairy Sci.* 95:3212-3217.

Shah, M. A. and M. R. Murphy. 2006. Development and evaluation of models to predict the feed intake of dairy cows in early lactation. *J. Dairy Sci.* 89:294-306.

Soriani, N., E. Trevisi, and L. Calamari. 2012. Relationships between rumination time, metabolic conditions, and health status in dairy cows during the transition period. *J Anim Sci* 90:4544-4554.

Soriani, N., G. Panella, and L. Calamari. 2013. Rumination time during the summer season and its relationships with metabolic conditions and milk production. *J. Dairy Sci.* 96:5082-5094.

St-Pierre, N. R. 2003. Reassessment of biases in predicted nitrogen flows to the duodenum by NRC 2001. *J. Dairy Sci.* 86:344-350.

Welch, J. G. and A. M. Smith. 1970. Forage quality and rumination time in cattle. *J. Dairy Sci.* 53:797-800.

**Chapitre 3. Valeur informative des durées de rumination
et de l'activité générale pour la détection des
troubles de santé**

Partie I. Intérêt du monitoring des baisses de durées de rumination ou d'activité générale pour la détection des troubles de santé chez les vaches laitières

Soumis. J Dairy Sci. 2014

1. Résumé de la Partie I

L'objectif de cette étude était d'évaluer la valeur informative d'alertes de santé produites à l'aide de tests « cumulative sum » (CUSUM), appliqués sur les données générées par un capteur de rumination et d'activité, pour la détection des principaux troubles de santé rencontrés en élevage.

Les données de rumination et d'activité ont été collectées dans 2 fermes expérimentales de l'ouest de la France sur 171 vaches Holstein durant 2,5 ans. Les alertes de santé fournies rétrospectivement par les tests CUSUM ont été évaluées en comparaison à la détection par observation visuelle par les éleveurs, qui était la méthode de référence. La nature, la sévérité et la date du trouble de santé étaient enregistrées de manière standardisée par les éleveurs. La sensibilité, la spécificité, le taux de faux positifs, les valeurs prédictives positives et négatives ont été calculées pour des paramétrages de CUSUM permettant une spécificité élevée (98% et 95% pour chacun des CUSUM sur l'activité et les durées de rumination) ou une sensibilité élevée (60% de la totalité des troubles de santé détectés par chacun des CUSUM). La précocité des alertes de santé par rapport à la détection par les éleveurs a également été évaluée.

Les troubles de santé les plus fréquents étaient les mammites (n = 172), les boiteries (n = 129), les atteintes de l'état général sans cause identifiée (n = 63) et les diarrhées (n = 19). Avec les paramétrages de CUSUM recommandés par les auteurs pour une utilisation en élevage (spécificité = 90%, taux de faux positifs = 4 détections/j pour 100 vaches présentes), 40% des troubles de santé détectés par les éleveurs ont donné lieu à des alertes par le CUSUM entre 4 j avant et 1 j après la détection par les éleveurs. La sensibilité des alertes CUSUM variait de 28% pour les boiteries modérées à 85% pour les mammites avec atteinte de l'état

général. Des sensibilités supérieures à 55% ont été obtenues pour les atteintes de l'état général sans cause identifiée, les diarrhées, et la plupart des troubles de santé graves mais plus rares. En raison de la spécificité imparfaite des alertes de santé et de la faible prévalence de troubles, la valeur prédictive positive des détections CUSUM était inférieure à 10% dans notre échantillon. Cependant, pour environ 50% des troubles de santé détectés l'alerte de santé avait lieu au moins 1 jour avant la détection par les éleveurs, permettant ainsi une alerte précoce.

Ces résultats suggèrent que, même si une grande variété de troubles de santé ont été effectivement détectés, les algorithmes utilisés dans cette étude seraient utiles uniquement en complément de l'observation visuelle pour la détection des troubles de santé par l'outil de monitoring.

2. Interpretative summary

By Clément P. et al., p xxx. Being able to efficiently monitor the health and the reproduction of their cows is a strong demand from dairy farmers. The aim of this study was to investigate the informative value for the detection of health disorders, of alarms provided by algorithms applied on rumination and activity data originating from a cow health and reproduction monitoring device. Our results indicate that, in addition to visual observation, the detection algorithms developed in this study may help farmers for early detection of a large variety of moderate to severe health disorders.

Interest of monitoring drops in rumination time and neck activity for the detection of health disorders in dairy cows

P. Clément^{*†}, R. Guatteo^{*†}, R. Quiniou[‡], N. Bareille^{*†}

^{*}LUNAM Université, Oniris, UMR BioEpAR, CS 40706, 44307 Nantes, France

[†]INRA, UMR1300 BioEpAR, CS 40706, 44307 Nantes, France

[‡] INRIA Rennes Bretagne - Atlantique, campus de Beaulieu 35042 Rennes Cedex

Corresponding author: pierre.clement@oniris-nantes.fr

3. Abstract

The objective of the current study was to evaluate the informative value of the detections provided for different health disorders (HD) in dairy cows by a one-sided cumulative sum (CUSUM) applied on data arising from a rumination and activity sensor. Rumination and activity data were collected on 171 Holstein cows for 2.5 years on 2 French research farms. The detections provided by the CUSUM tests were evaluated in comparison with HD detection by farm staff, which was the reference method. Sensitivity, specificity, false detection rate, positive predictive value and negative predictive value were calculated for CUSUM settings enabling a high specificity (98% and 95% for CUSUM on rumination time or on neck activity) or a high sensitivity (60% of all HD detected by individual CUSUM). The precocity of the detections was also evaluated. The most frequent HD were mastitis (n = 172), lameness (n = 129), impaired general condition without identified cause (n = 63) and diarrhea (n = 19). With the settings recommended by the authors for an on-farm use (specificity = 90%, false detection rate = 4 detections/d for 100 cows), on average 40% of the HD detected by farm staff were detected by the CUSUM between 4 d before and 1 d after the detection by farm staff, with sensitivities varying from 28% for middle intensity lameness to 85% for mastitis with general clinical signs. Sensitivities above 55% were obtained for impaired general condition episodes without identified cause, diarrhea, and most of the rare but severe HD. Due to the lack of specificity of the detections and to the low prevalence of HD, the positive predictive values of CUSUM detections remained below 10% in our sample. However, approximately 50% of the detected HD were detected at least 1 d before the detection by farm staff, enabling early attention by farmers. Results suggest that, even if a large variety of HD were detected, the algorithms used in this study would be useful only in addition to visual appraisal for HD detection by the monitoring device.

Key words: health monitoring, dairy cow, sensor, algorithm

4. Introduction

In agriculture, there has been a general trend for decades towards automation of heavy work or time-consuming tasks. In the mid-1980s, the first Automated Milking Systems (AMS) were designed and, by 2009, more than 8,000 farms were equipped with AMS, mainly in western Europe (de Koning, 2010). However, this development has required the design of sensors and electronic devices to ensure the autonomy of AMS. One of the main challenges was to enable automated detection of mastitis, as abnormal milk must not be put on the market. Therefore, research focused initially on the development of sensitive mastitis detection devices, based on automated measures of milk electrical conductivity, color, or temperature (Hogeveen et al., 2010). Recently, tools have been developed to assist dairy farmers in cow health or reproduction management, for instance for estrus or calving detection (Bewley, 2009, Rutten et al., 2013). Different kinds of sensors measuring for instance activity, milk conductivity, temperature, ruminal pH, rumination time or feeding time have been used (Bewley, 2009) to detect cows with deviating values, which deserve special attention from the farmer.

One of the characteristics of monitoring devices is the spectrum of events which they can detect. Some devices are mostly dedicated to the detection of one type of event, for instance mastitis, estrus or lameness (Rutten et al., 2013), whereas others use sensors that measure a parameter potentially impacted by various reproduction or health disorders (temperature sensors (Adams et al., 2013), rumination or activity sensors (Chapinal et al., 2010)). In order to enlarge the spectrum of events that are susceptible to be detected, some devices combine the data provided by 2 or more sensors to trigger health or reproduction alarms. For such “broad spectrum” tools, it is crucial to evaluate which events are likely to be detected and what is the performance of this detection. Indeed, a lack of sensitivity or specificity of the detections may (i) alter farmers’ confidence in the tool and may lead to its underuse (ii) impact farm profitability as non-optimal disease management results in economic losses due to decreased milk yield, increased treatment costs or increased risk of culling (Fourichon et al., 1999, Seegers et al., 2003) (iii) compromise animal welfare if some cows are not detected as diseased or are detected late (EFSA, 2009) (iv) promote excessive use of antibiotic treatments in case of an excessive number of false alarms. However, the performance of the commercial devices in terms of detection quality is very often not given by the manufacturers.

Recently, a rumination and activity sensor (HR-Tag, SCR Engineers Ltd., Netanya, Israel) has been developed. This sensor is supported by a neck collar and continuously records rumination time (**RT**) and neck activity (**NA**) of each cow of the herd. Sensor data are used to generate lists of cows that need special attention concerning estrus (overactivity) or health (drops in activity or rumination time, or both). Performance of this automated system for estrus detection has been evaluated elsewhere (Kamphuis et al., 2012). Some health disorders (HD) such as mastitis (Fitzpatrick et al., 2013) or lameness (Van Hertem et al., 2013) have been shown to provoke drops in RT or NA which are detectable with this sensor. However, these studies have concerned disease occurrences in experimental conditions on very few animals, or the detection of one single type of HD. Other studies focused on describing the relationship between the RT measured by this sensor and the metabolic or health status during the transition period (Soriani et al., 2012) or the summer season (Soriani et al., 2013a). These studies described statistical associations between RT and health status without an objective of being able to detect the cows with impaired health or metabolic status. In practice, RT and NA are indicators that may be affected in a large number of HD (Radostits et al., 2007) and therefore they could be used to detect diseased cows. Thus, we can assume that all HD, depending on their type and severity, will impact the parameters measured by this sensor with different magnitudes.

The aim of this study was to evaluate the informative value of a one-sided cumulative sum (CUSUM) (Luo et al., 2009) applied on RT and NA measured by HR-Tag, in detecting significant shifts related to different health disorders.

5. Materials and methods

Cows and Herd Management

The data used in this study were collected on 2 research farms in western France where management practices were representative of classical farming conditions in this area. Cows were milked twice daily and fed a corn silage based TMR or a pasture based diet, depending on the period of the year. On farm 1, (Les Trinottières: Maine-et-Loire, France)

approximately 115 Holstein cows were involved between September 01, 2010 and August 31, 2013. Cows were managed in 2 batches: 1 (60 cows) fed a TMR containing 75% corn silage and 25% concentrate (zero-grazing) and the other batch fed the same TMR from November to April and fed a diet containing 50% grass and 50% corn silage from April to November. On farm 2, (Trevarez: Brittany, France), 56 Holstein cows were involved between November 04, 2009 and January 11, 2012. Cows were fed a corn silage based TMR during winter or a pasture based diet from spring to autumn. Cows were housed in cubicles during winter on both farms, and average milk yield was 30.6 kg/d (farm 1) and 25.9 kg/d (farm 2).

Health Disorder Recording

Throughout the study, all cows were monitored daily by farm staff for signs of disease according to standardized protocols. Farm staff had no access to sensor data or CUSUM detections during the study. Thus, RT or NA could not be used to identify cows with HD. Nature, severity and date of occurrence of HD detected were recorded by farm staff on a standardized form (table 3-I-1). On each farm, one person was in charge of recording HD and the authors ensured homogenous recording between farms.

Table 3-I-1: standard form used by farm staff for recording health disorders. Definition of the type and severity of the health disorders

	Severity	Description
Mastitis	1	Clots in milk only
	2	Clots in milk, udder inflammation, red and painful quarter
	3	Mastitis with general clinical signs: hyperthermia or drop in milk or decreased appetite
Lameness	1	Preventive hoof treatment
	2	The cow is obviously lame on 1 or more legs. An observer will in most cases be able to tell which leg is affected.
	3	The cow is obviously lame on 1 or more legs. The cow is unable, unwilling, or very reluctant to bear weight on the affected leg.
Milk fever	1	Treatment of a cow suspected to begin a milk fever
	2	Treatment of a cow down with milk fever symptoms
Retained placenta		Retained placenta more than 24 h post-partum with general clinical signs: hyperthermia or drop in milk or decreased appetite
Metritis	1	Vulvar muco-purulent material discharge before 21 d postpartum
	2	Endometritis: vulvar discharge with purulent material after 21 d postpartum, without alteration of general condition
	3	Acute metritis: vulvar discharge with purulent material before 21 d postpartum with impaired general condition: hyperthermia or drop in milk or decreased appetite

Impaired general condition	1	Impaired general condition without identified cause: drop in milk or decreased appetite without hyperthermia.
	2	Impaired general condition without identified cause: drop in milk or decreased appetite with hyperthermia.
Other health disorder		Description as precise as possible of the health disorder

Sensor Data

At calving or a few days before, depending on standard procedures in the farm, cows were fitted with a neck collar supporting a NA and RT logger (HR-Tag, SCR Engineers Ltd., Netanya, Israel). The logger contains a microphone that records RT continuously and summarizes it into 2 h blocks (Schirmann et al., 2009). It also contains an accelerometer that, after algorithmic treatment, summarizes NA into 2 h blocks (Kamphuis et al., 2012). RT and NA data were stored in the logger and transmitted to a computer (DataFlow software, SCR Engineers Ltd., Netanya, Israel) when cows walked below an infrared antenna located in the milking parlor or over a trough. Thus, new data were typically available twice a day after milking and data were often not available during dry period.

Algorithms for the Detection of Drops in RT and NA

As the commercial algorithms developed by the manufacturer for HD detection with the device used in this study were confidential, the authors of the present study developed their own algorithms. The algorithms were retrospectively applied on sensor data.

A method based on a one-sided cumulative sum (CUSUM) test (Luo et al., 2009) was used to detect drops in RT or NA. The CUSUM test accumulates the difference (error) between the reference RT or NA and the observed RT or NA if this difference is positive (figure 3-I-1).

First, a reference RT and NA was calculated for each cow and for each 2 h block of the day (12 blocks/d) as the mean RT and NA respectively, recorded for this cow at the same 2 h time block of the day during the past 7 d. This way of calculation enables to take into account circadian variations of RT and NA. Values equal to 0, often due to technical problems on the sensors, were considered as missing values and were not included in the reference calculation. The computation of the reference began as soon as the cows were fitted with the sensors. During the first 7 d, it was computed with the available RT and NA data from an incomplete period. Secondly, an alarm was triggered when the CUSUM value exceeded a given decision interval (h). The CUSUM used in this study was computed for each cow as:

$$\begin{cases} C_0 = 0 \\ C_t = \max(0 ; C_{t-1} + \varepsilon_t - k) \end{cases}$$

Where C_t is the CUSUM value at time t , ε_t is the difference between the reference RT or NA values and the observed values at time t , k is equal to $b \times \sigma_{(\text{ref-obs})7d}$, where b is a constant and $\sigma_{(\text{ref-obs})7d}$ is the standard deviation of the error (difference between observed and reference values) calculated for RT or NA for each cow on all the values of the last 7 d, whatever the time block of the day. This last parameter was added to take into account the high variability of RT and NA between cows. In order to limit the number of very short alarms, often corresponding to irrelevant events, a detection was triggered only if the CUSUM exceeded the decision interval (h) during at least 2 x 2 h blocks. The CUSUM value was reset to 0 if more than 12 successive values of RT or of NA were missing.

In a first step, combinations of (b, h) values were tested with b varying from 0.3 to 0.8 with a 0.1 unit increase each time. For each b value, h values between 10 and 100 were tested with a 10 unit increase each time. The final b value for the CUSUM on RT and NA data was set to 0.7, as this value gave the highest level of specificity for a given level of sensitivity, and gave the highest level of sensitivity for a given level of specificity. In a second step, different decision intervals (h) were applied on RT data or NA data to reach either a specificity of 98%, a specificity of 95%, or a sensitivity of 60% (see below for sensitivity and specificity calculations) for the HD occurring more than 5d before calving or after 11 DIM (**HD**). These quite high values of specificity and sensitivity were chosen because, on the one hand, a decrease in specificity quickly induces a high proportion of false positive detections, but on the other hand a high sensitivity is desired by the farmers. CUSUM algorithm's specificity

was evaluated for a given sensitivity and vice versa. Depending on the type of HD, RT and NA may not be affected with the same magnitude. Then, separate analyses were carried out for the CUSUM on RT data and for the CUSUM on NA. In a final step, the added value of combining both the CUSUM on RT and on NA data (combined algorithm) was evaluated. First, the CUSUM on RT data and the CUSUM on NA data were applied separately on data with settings enabling to reach a given level of specificity (98 or 95%) or a given level of sensitivity (60%) for each CUSUM. Then, for each level of specificity or of sensitivity of individual CUSUM, the settings found in step 1 were used and CUSUM on RT and on NA were applied together on data. Finally, the informative value of a detection by the combined algorithm (CUSUM on RT or CUSUM on NA) was calculated.

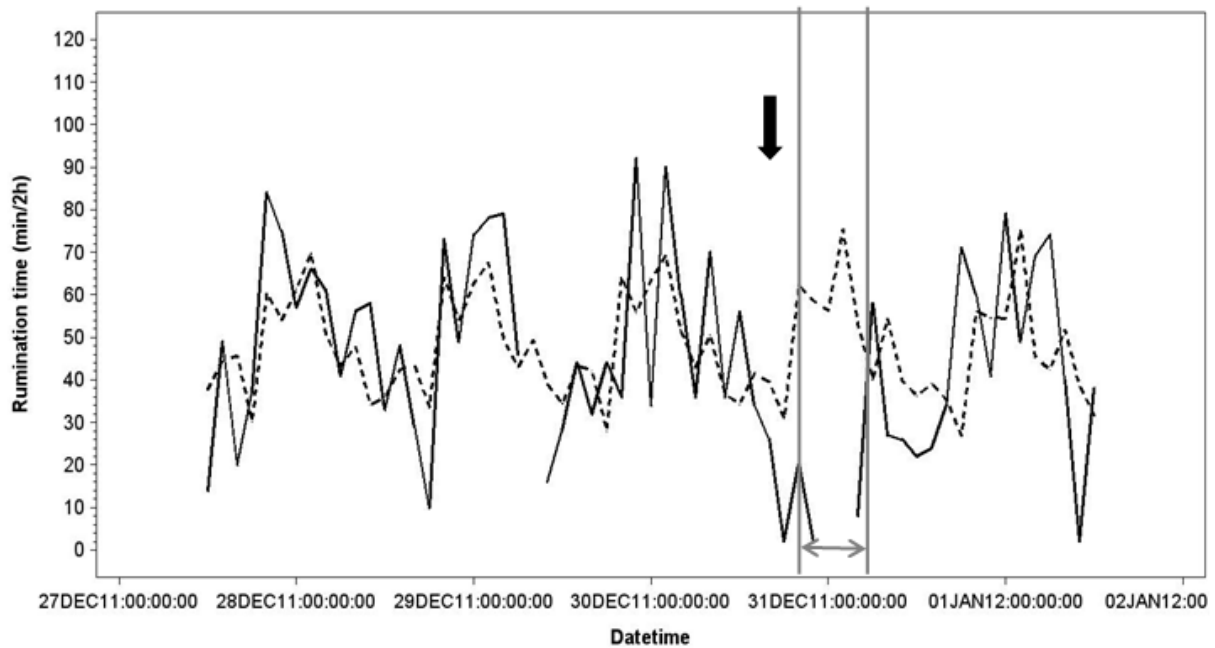


Figure 3-I-1. Example of rumination time measured for one cow around an health disorder (solid line). The reference rumination time (dotted line), computed as the mean rumination time measured for that cow on the same 2 h time block of the day during the past 7 d is also shown. One significant difference (horizontal arrow) between measured and reference rumination time was identified just after a mastitis detection (vertical arrow), using a method based on the cumulative sum test.

Informative Value of the Detections

As a monitoring tool may detect an HD a few hours or a few days before or after the detection by farm staff, the definition of a time-window in which a detection by the tool is likely to be caused by the HD is necessary for informative value calculation (Hogeveen et al., 2010). The exact time of detection of HD was not known in our data set, so the detection by farm staff was arbitrarily positioned at 1200 h on the day of detection. For sensitivity calculation, a 5 d time-window starting 96 h (4 d) before the time of the detection of the HD by farm staff and ending 24 h (1 d) after was defined (figure 3-I-2). This 5 d time-window was called the validation time-window. The 5 d timeframe was chosen because too long a timeframe may lead to the consideration of detections not linked to a given HD – i.e. false positives - as true positives and because receiving an alarm 1 week before the occurrence of putative clinical symptoms is of low interest for the farmer. On the contrary, too short a timeframe may lead to the consideration of some truly positive detections as false positives because of too short a validation time-window. If 2 or more HD were close enough in time to provoke time-window overlap, only one validation time-window starting 4 d before the first HD and ending 1 d after the last HD was considered for analysis. A validation time-window was excluded from analysis if more than 33% of RT or NA were missing in this time-window. For each CUSUM setting of interest, a HD was considered as a True Positive (**TP**) if a detection was triggered during the validation time-window containing this HD (figure 3-I-1). An HD was considered as a False Negative (**FN**) if no detection was triggered during the validation time-window containing this HD. The sensitivity of the CUSUM detections was calculated for each type and severity of HD as $Se = TP / (TP + FN)$.

Calving has been shown to provoke a decrease in RT (Schirmann et al., 2013) and an increase in activity (Jensen, 2012). Thus, for HD occurring around calving, the time-lag between calving date and HD detection had to be taken into account for specificity calculation. Distinct specificity calculations were carried out for the peripartum period and for the lactation or dry period. Concerning the period prior to 5 d before calving or after 11 DIM (lactation-dry period), three 5 d control time-windows were randomly sampled (proc surveyselect, SAS 9.2) for each cow with sufficient sensor data (more than 3.5 months of data). To be validated, these 5 d control time-windows had to start after 16 DIM, to start at least 20 d before or 20 d after an HD, to be distant from another 5 d control time-window by a minimum of 20 d, and

to contain less than 33% of missing RT or NA data. For each CUSUM setting of interest, a control time-window was considered as a False Positive (**FP**) if a detection was triggered during this time-window. A control time-window was considered as a True Negative (**TN**) if no detection was triggered during this time-window. The specificity of the CUSUM detections during lactation and dry period (**Sp**) was calculated as $Sp = TN / (TN + FP)$. Among the 222 cows of the study with more than 3.5 months of data, 213 could be correctly sampled due to the constraints imposed to validate the control time-windows. Thus, specificity calculations were made on 639 (3 x 213) 5 d control time-windows.

For HD occurring between 5 d before and 11 d after calving (peripartum period), the specificity of CUSUM detections (**Sp_{pp}**) was calculated as follows. Firstly, each HD occurring in the peripartum period, characterized by its time-lag to calving, was matched with three 2 h blocks of data with the same time-lag to calving, randomly sampled among cows that did not have a HD between 5 d before and 11 d after calving. Secondly, for each date and time sampled, a control time-window beginning 4 d before and ending 1 d after the date and time sampled was created. To be validated, these 5 d control time-windows had to contain less than 33% of missing RT or NA data. Then the CUSUM algorithms were applied on control time-windows and the specificity of the CUSUM algorithms was calculated as explained above. The specificity was calculated on 219 (3 x 73) control time-windows.

For sensitivity and specificity calculations, an exception was made for detections occurring for the CUSUM on RT data between 8 h before and 26 h (estrus time-window) after the beginning of an overactivity episode detected by the commercial algorithm of estrus detection associated with the device (Kamphuis et al., 2012). As a decrease in RT is often observed around estrus (Reith and Hoy, 2012), these drops in RT detected in the estrus time-window around activity peaks (overactivity algorithm value > 5.2) were neither counted as TP (if in a validation time-window) nor as a FP (if in a control time-window) but instead were ignored.

In order to propose a user-transparent measure of specificity, an indicator linked to specificity was calculated as follows. For each cow and for each day with sufficient sensor data (i.e. with less than 6/12 RT or NA missing values) outside the validation time-windows, a dichotomous variable (**Y**) was created. For one day *d*, the value **Y**(*d*) was set to 0 if no detection was triggered by the CUSUM between 0000 h and 2359 h, and was set to 1 if at least 1 detection was triggered by the CUSUM during the same time-window. In case of long-lasting detections or repeated detections for a cow, the variable was set to 1 for every day in which

the cow was triggering a detection. Then the mean number of cows triggering a detection/day was calculated for different settings of the CUSUM as false detection rate = (number of days-cows where $Y = 1$)/(number of cow days where $Y = 1$ or $Y = 0$) x 100. It represents, for 100 cows present in the herd, the mean number of detections triggered daily by the CUSUM outside validation time-windows i.e. when no HD were detected by farm staff.

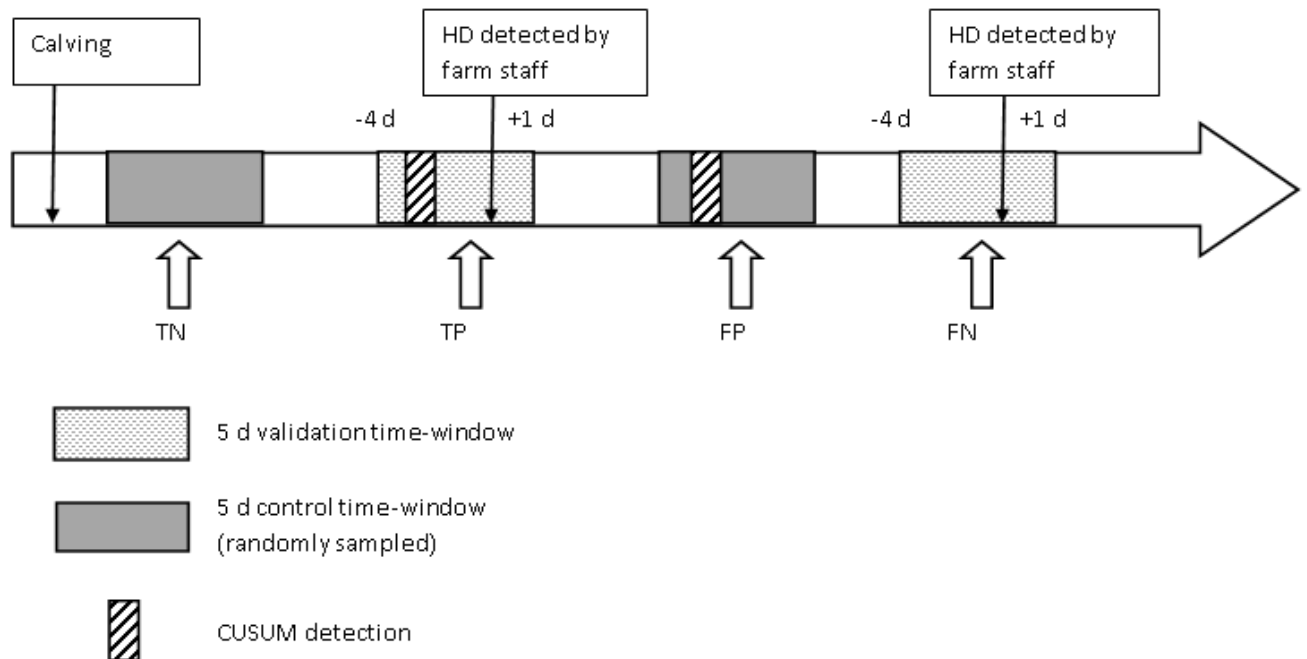


Figure 3-I-2. Principle of computation of events for sensitivity and specificity calculation (TN = True Negative. TP = True Positive. FP = False Positive. FN = False Negative).

Predictive Values of CUSUM Detections

Predictive values are operational values of major importance to evaluate the usability of a detection tool. Positive Predictive Values and Negative Predictive Values during lactation or dry period were considered being dependent on the proportion of diseased cow days of the study, calculated as follows. Firstly, the total number of cow days of the trial with less than 6/12 RT or NA missing values was calculated. Secondly, the number of cow days where it was relevant to detect an HD over trial, called diseased cow days, was calculated. For that

purpose, each HD was considered to last for 3 d and the number of diseased cow days was equal to 3 x (number of HD). This 3 d duration was chosen because it is of interest to detect diseased cows during a relatively short period of time around HD and because more than 66 % of the detection by the CUSUM on RT or NA data occurred between -1 d and +1 d compared to farm staff detection. Thirdly, the number of healthy cow days of the trial was computed by subtracting the number of diseased cow days from the total number of cow days. Fourthly, a theoretical number of TP, TN, FP and FN (TP_{th} , TN_{th} , FP_{th} and FN_{th}) was calculated knowing the sensitivity and the specificity of the CUSUM algorithms. For example $TP_{th} = Se \times (\text{number of diseased cow days})$ and $FN_{th} = (\text{number of diseased cow days}) - TP_{th}$. Fifthly, the Positive Predictive Value was calculated as $PPV = TP_{th}/(TP_{th} + FP_{th})$ and the Negative Predictive Value was calculated as $NPV = TN_{th}/(TN_{th} + FN_{th})$.

Finally, PPV and NPV were calculated following the same principle for a range of prevalence, from 0.5% of diseased cow days to 5% of diseased cow days (0.95% of diseased cow days in our sample) to evaluate the interest of CUSUM algorithms in different HD prevalence conditions. Calculations were made for the detection by the CUSUM on RT data or on NA data (combined algorithm).

Time-lag between CUSUM Detections and Farmer's Detections

Another important aspect when evaluating a monitoring tool is the putative earlier detection of some health disorders enabled by the use of that tool, compared to farmer detection. Indeed, an early detection may lead to earlier treatment with improved cure rates. To investigate this putative gain in terms of precocity of the detection, only validation time-windows with 1 HD actually detected were included in the analysis to be sure to attribute a detection to 1 health event. For each type of HD with more than 20 events detected, the time-lag between the first detection by the CUSUM algorithm on RT data or on NA data in the validation time-window and the detection by farm staff was calculated. Calculations were made for the detection by the CUSUM on RT or on NA data with the settings enabling to reach a specificity of 95% for each CUSUM algorithm.

6. Results

Descriptive Statistics of the HD Detected

Initially, 668 HD were detected (table 3-I-2). Among these, the HD that were not known in the literature to provoke sudden drops in rumination time or in activity (lameness severity 1, metritis severity 1 or 2) were ignored (n = 116). For accidents (n = 8), defined as rare or low severity individual events that occurred during the study (e.g. cows involuntarily blocked for 1 d in their stalls, or a cow with epistaxis for a few hours), the entire 5 d time-window around the event (-4 d to +1 d) was excluded from the analyses for both sensitivity and specificity calculations. The same procedure was applied for dystocia or caesarian section events (n = 5). In a second step, 25% of the HD were excluded from the analysis due to excessive RT or NA missing values (more than 33% of missing values). Finally, 404 HD were included in the analysis, with 348 HD detected on farm 1 and 56 HD detected on farm 2. The most frequent HD were mastitis (n = 172), lameness (n = 129), impaired general condition without identified cause (n = 63) and diarrhea (n = 19). Other types of HD were very rare. Within a type of HD, the number of HD decreased with HD severity. Overall, among the 259 cows of the data set, 170 cows had at least 1 HD during the study. The mean number of HD per cow was 1.6 (range 0-14).

Table 3-I-2. Number of health disorders detected by farm staff over the study (Ig=Ignored for the analysis, E=Excluded from the analysis)

Health disorders	Severity	Initial	Eligible for analysis	Included in the analysis ¹		
				Farm1	Farm2	Total
Mastitis	1	144	144	90	21	111
	2	71	71	30	18	48
	3	19	19	11	2	13
Total Mastitis		234	234	131	41	172
Lameness	1	63	Ig	Ig	Ig	Ig
	2	115	115	84	10	94
	3	40	40	35	0	35
Total Lameness		218	155	119	10	129
Impaired general condition	1	46	46	38	1	39
	2	26	26	24	0	24
Total impaired general condition		72	72	62	1	63
Digestive tract disorders						
Diarrhea		26	26	16	3	19
Colics		1	1	1	0	1
Left displaced abomasum		3	3	1	0	1

Others						
Down after a fall on concrete		6	6	5	0	5
	1	22	Ig	Ig	Ig	Ig
Metritis	2	31	Ig	Ig	Ig	Ig
	3	4	4	3	0	3
Total Metritis		57	4	3	0	3
Pericarditis		1	1	1	0	1
Pneumonia		2	2	2	0	2
Abortion		3	3	2	0	2
Dystocia or caesarian section		5	E	E	E	E
Retained placenta		7	7	3	0	3
Milk fever	1	6	6	1	0	1
	2	19	19	1	1	2
Total milk fever		25	25	2	1	3
Accident		8	E	E	E	E
Total		668	539	348	56	404

Informative Value of CUSUM Detections

The informative value of CUSUM detections applied on RT or NA data is displayed in table 3-I-3. The column on the left of the table shows the results obtained with CUSUM settings enabling a 98% specificity (CUSUM settings 1), the middle column block shows the results obtained with CUSUM settings enabling a 95% specificity (CUSUM settings 2), and the right column block shows the results obtained with CUSUM settings enabling a 60% sensitivity for individual CUSUM algorithms during lactation and dry period (CUSUM settings 3). For each CUSUM setting, the specificity of CUSUM detections during lactation and dry period, during the peripartum period, the false detection rate and the global sensitivity of CUSUM detections, i.e. when regrouping HD and HD_{pp}, are displayed in the upper part of the table. With CUSUM setting 1, the CUSUM on RT, on NA and on RT/NA (= RT or NA, combined algorithm) allowed the detection of 15.6, 13.4 and 21.3% of the HD respectively. The false detection rates were 0.8, 0.8 and 1.6 detections/d respectively for 100 cows. The combination of the CUSUM on RT or on NA induced a small decrease in specificity (96.7% vs. 98.0%). The specificity obtained for peripartum HD was clearly lower than the specificity calculated during lactation and dry period (-3.5 to -15.4%). With CUSUM setting 2, 23.3 to 39.6% of the HD were detected with the false detection rate varying from 1.5 to 4 detections/d for 100 cows. The combination of CUSUM algorithms on RT or on NA induced a sharp decrease in specificity (90.1% vs. 95.0%). With CUSUM setting 3, 57.4 to 76.5% of the HD were detected but the value of the false detection rate was very high (9.0 to 19.0 detections/d for 100 cows) due to a very low specificity (26.9% to 53.0%). In practice, such a high rate of false positive detections greatly limits the interest of the information given by the device. Thus, only the results obtained with CUSUM settings enabling a high specificity (settings 1 and 2) will be described in details below.

The lower part of table 3-I-1 shows the sensitivity obtained for each HD type and severity for the different CUSUM settings. With a given CUSUM setting, the combined algorithm (CUSUM on RT/NA) was always much more sensitive than the CUSUM based solely on RT or on NA. That means that a large number of HD that were detected by the CUSUM on RT were not detected by the CUSUM on NA and vice-versa. There was no tendency for one of the 2 CUSUM (on RT or on NA) to be systematically more sensitive for the detection of one type of HD: for instance, with CUSUM setting 1, more impaired general condition episodes

without hyperthermia were detected by the CUSUM on RT (Se = 46.2% for the CUSUM on RT vs. 28.2% for the CUSUM on NA) while an opposite result was obtained for impaired general condition episodes with hyperthermia. Even lameness episodes were not always detected with a clearly better sensitivity by the CUSUM on NA, especially with CUSUM setting 1. As expected, the greater the severity of HD, the better the sensitivity of the CUSUM detections. On average, with CUSUM setting 1, 15.7% (14.4-46.2% depending on severity) of mastitis, 11.6% (7.4-22.9%) of lameness, 49.2% (45.8-51.3) of impaired general condition and 36.8% of diarrhea were detected by the CUSUM on RT or NA. With CUSUM setting 2, 34.9% (29.7-84.6%) of mastitis, 30.2% (27.7-37.1%) of lameness, 63.5% (61.5-66.7%) of impaired general condition and 57.9% of diarrhea were detected. Quite rare but severe events such as colic, acute metritis, left displaced abomasum or pericarditis were generally detected with CUSUM setting 2 (data not shown). Neither of the 2 abortions of the data set was detected with these settings.

Table 3-I-3: Informative values of CUSUM detections applied on Rumination Time (RT), Neck Activity (NA) or (RT or NA) for 3 CUSUM settings, for the main health disorders detected by farm staff

CUSUM decision limits	CUSUM settings 1			CUSUM settings 2			CUSUM settings 3			
	$h_{RT} = 100$	$h_{NA} = 75$	$h_{RT} = 100$ $h_{NA} = 75$	$h_{RT} = 75$	$h_{NA} = 36$	$h_{RT} = 75$ $h_{NA} = 36$	$h_{RT} = 30$	$h_{NA} = 13$	$h_{RT} = 30$ $h_{NA} = 13$	
Sp/Se (%)	Sp=98.0		Sp = 96.7	Sp = 95.0		Sp = 90.1	Se = 60.0		Se = 76.8	
CUSUM	RT	NA	RT or NA	RT	NA	RT or NA	RT	NA	RT or NA	
Sp (%)	98.0	98.0	96.7	95.0	95.0	90.1	69.0	69.0	50.9	
Sp _{PP} ¹ (%)	85.4	94.5	81.3	81.3	80.4	66.2	53.0	46.6	26.9	
FDR ²	0.8	0.8	1.6	1.5	2.8	4.0	9.0	12.0	19.0	
Global Se (%) (n = 404)	15.6	13.4	21.3	23.3	29.5	39.6	57.4	60.4	76.5	
HD	Severity	Se by HD (%)			Se by HD (%)			Se by HD (%)		
Mastitis	1 (n = 111)	9.0	6.3	14.4	14.4	19.8	29.7	49.5	48.6	68.5
	2 (n = 48)	10.4	6.3	10.4	16.7	25.0	33.3	62.5	52.1	79.2
	3 (n = 13)	30.8	38.5	46.2	53.8	61.5	84.6	76.9	92.3	100.0
	Total (n = 172)	11.0	8.7	15.7	18.0	24.4	34.9	55.2	52.9	73.8
Lameness	2 (n = 94)	4.3	4.3	7.4	9.6	21.3	27.7	43.6	53.2	68.1
	3 (n = 35)	11.4	17.1	22.9	20.0	31.4	37.1	65.2	74.3	91.4
	Total (n = 129)	6.2	7.8	11.6	12.4	24.0	30.2	51.2	58.9	74.4
Impaired general condition	1 (n = 39)	46.2	28.2	51.3	53.8	43.6	61.5	66.7	69.2	79.5
	2 (n = 24)	33.3	45.8	45.8	45.8	62.5	66.7	75.0	87.5	87.5
	Total (n = 63)	41.3	34.9	49.2	50.8	50.8	63.5	69.8	76.2	82.5
Diarrhea	(n = 19)	31.6	21.1	36.8	47.4	31.6	57.9	63.2	73.7	84.2
Down after a fall on concrete	(n = 5)	20.0	40.0	40.0	20.0	60.0	60.0	80.0	100.0	100.0
Metritis	3 (n = 3)	0.0	33.3	33.3	0.0	33.3	33.3	100.0	66.7	100.0
Retained placenta	(n = 3)	0.0	0.0	0.0	0.0	0.0	0.0	66.7	33.3	100.0
Milk fever	1 (n = 1)	0.0	0.0	0.0	0.0	0.0	0.0	0.0	100.0	100.0
	2 (n = 2)	100.0	0.0	100.0	100.0	50.0	100.0	100.0	100.0	100.0
	Total (n = 3)	66.7	0.0	66.7	66.7	33.3	66.7	66.7	100.0	100.0

¹ Specificity for the peripartum period

² False detection rate, expressed in detection/d for 100 cows

The PPV and NPV obtained for the CUSUM algorithms on RT/NA for the different CUSUM settings are displayed in table 3-I-4. With the prevalence of diseased cow days observed in our study sample (0.95%), the PPV of a detection by the CUSUM on RT/NA obtained with CUSUM settings 1 and 2 were 5.9 and 3.7% respectively and the NPV of a detection were 99.2 and 99.3% respectively. With prevalence ranging from 0.5 to 5%, the PPV varied from 3.1 to 24.2% with CUSUM setting 1, and from 1.9 to 16.2% with CUSUM setting 2. Concomitantly, NPV remained high, always above 96%.

Table 3-I-4. Predictive values of CUSUM on RT or NA detections for different HD prevalences (PPV = Positive Predictive Value, NPV = Negative Predictive Value)

Prevalence¹ (%)	CUSUM settings 1 (Sp = 96.7%) h _{RT} = 100 h _{NA} = 75	CUSUM settings 2 (Sp = 90.1%) h _{RT} = 75 h _{NA} = 36	CUSUM settings 3 (Sp = 50.9%) h _{RT} = 30 h _{NA} = 13
0.5	PPV = 3.1% NPV = 99.6%	PPV = 1.9% NPV = 99.7%	PPV = 0.8% NPV = 99.8%
0.9	PPV = 5.9% NPV = 99.2%	PPV = 3.7% NPV = 99.3%	PPV = 1.5% NPV = 99.6%
2	PPV = 11.3% NPV = 98.4%	PPV = 7.2% NPV = 98.7%	PPV = 3.0% NPV = 99.1%
5	PPV = 24.2% NPV = 96.1%	PPV = 16.2% NPV = 96.7%	PPV = 7.2% NPV = 97.8%

¹ Prevalence expressed as the percentage of diseased cow days (= 3 x number of HD) among the total number of cow days with sufficient sensors data over the trial. Prevalence in the study sample = 0.9%

For HD with more than 20 events detected before 5 d prior to calving or after 11 DIM, the time-lag between farm staff detections and CUSUM on RT/NA detections for CUSUM settings 2 is displayed in figure 3-I-3. For the 3 types of events (mastitis, lameness, impaired general condition), around 50% of the HD were detected at least one day before the detection by farm staff with 4 to 15% detected 3 days before. Thirty percent of the events were detected on the same day as farm staff and 20% were detected 1 day after.

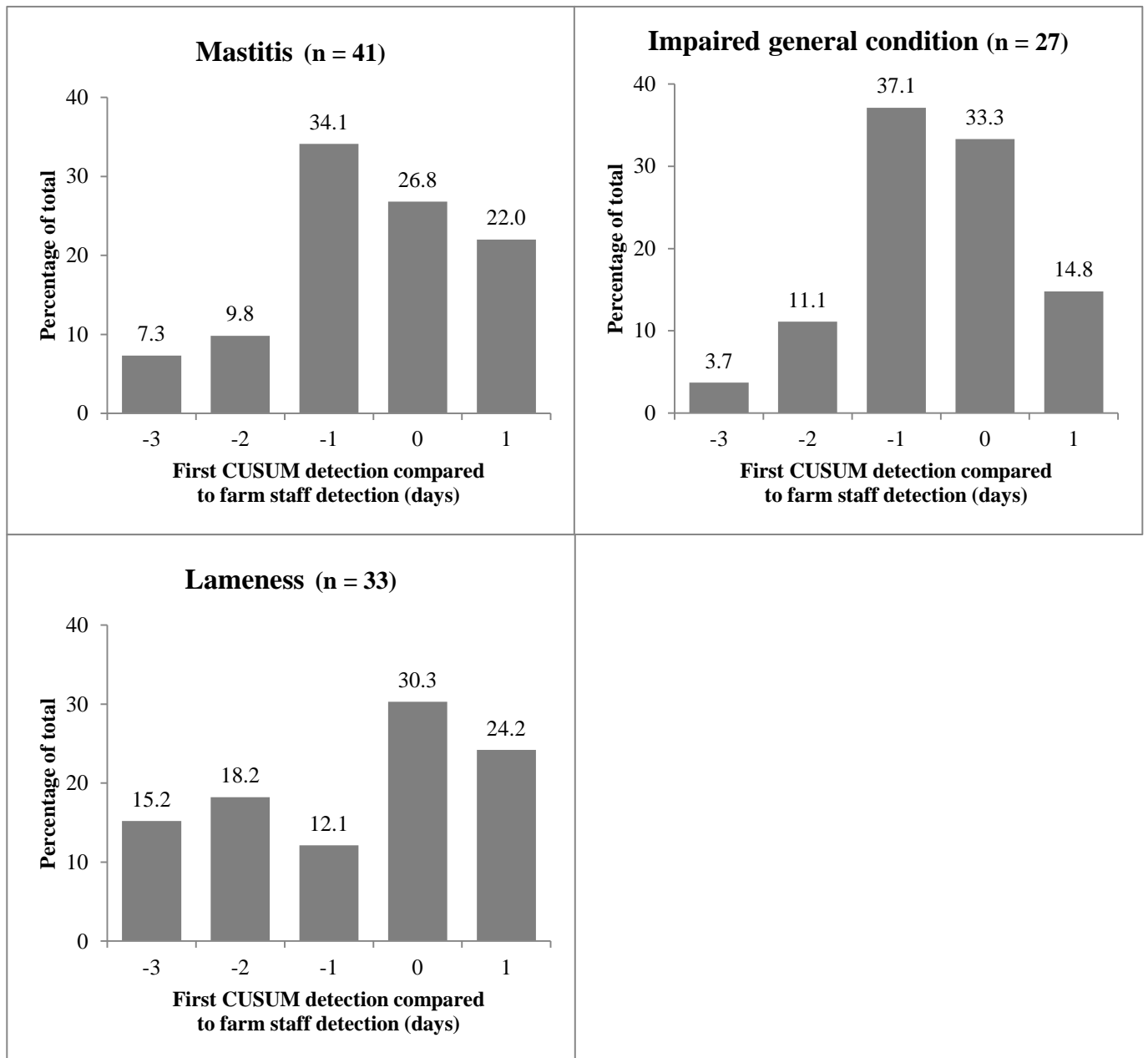


Figure 3-I-3. Distribution of the time of first detection by the CUSUM on RT or on NA compared to the detection by farm staff for the main health disorders.

7. Discussion

Our study is the first one to evaluate the informative value of CUSUM algorithms applied on raw RT and NA data provided by HR-Tag for the detection of a large variety of HD. In our opinion, CUSUM settings 2 should be used in practice, as it gives the best trade-off between global sensitivity of the detections (39.6%), the PPV (3.7%) and the false detection rate (4 detections/d for 100 cows). With these settings, the sensitivity ranged from 27.7 to 84.6%, depending on the type and severity of the HD considered. Mastitis with general clinical signs, impaired general conditions without identified cause, diarrhea and most of the rare but severe HD were detected with a sensitivity above 55%. This makes our algorithms interesting to detect the more severe HD in herds. However, with the moderate level of sensitivity reached, the monitoring device must be seen as complementary to visual appraisal by the farmer for HD detection. This result is not surprising as in a preliminary analysis of the literature, we failed to precisely determine what type of HD would provoke a sharp decrease in RT or NA, detectable by a CUSUM test. Thus, some HD included in the analysis such as local mastitis rarely provoked a detectable drop in RT or NA. Moreover, one CUSUM setting is optimal only for the detection of a given shift in RT or NA (Luo et al., 2009). In our study, the expected shift in RT or NA was unknown but was expected to be variable between and within HD types, placing the CUSUM test in non-optimal conditions of detection for some HD events. The same problem would have occurred with other detection methods. Another difficulty relative to algorithm development in this study was that raw RT and NA data were highly variable between cows and over time within cows. The variability of raw data, and particularly RT, was very high over time (figure 3-I-1). Thus, the development of efficient algorithms might be more difficult when working on highly variable time-series. We also chose a relatively short validation time-window compared with other researchers who used time-windows up to 17d for mastitis detection for instance (Hogeveen et al., 2010). This may partly explain the low sensitivities found. It is unlikely that an automated detection 17d before a clinical mastitis would be really linked to that mastitis episode.

Our results underline the interest of combining 2 measures (RT and NA) for the detection of HD. This is consistent with the results of other studies that demonstrated the interest of combining data from different sensors and milk yield recording systems for the detection of mastitis or lameness for instance (Kamphuis et al., 2008, Van Hertem et al., 2013). Similarly, for Rutten et al. (2013), the integration of non-sensor data (cow history), economic information or information coming from the farmer is clearly a way to improve sensor-based advice for health and reproduction management on farms. Many authors evaluated the informative value of the devices

used in our study, or very similar ones that measure only activity, for estrus detection. They found that the performance for estrus detection by the device was similar to the detection by visual appraisal (Holman et al., 2011) or slightly worse (Kamphuis et al., 2012), and better than the detection by one pedometer (Holman et al., 2011). Van Hertem et al. (2013) evaluated the informative values of a model that included rumination and activity data coming from the same sensors for lameness detection. These sensors' data were combined in the model with milk yield data. They found better informative values than in our study, with a sensitivity of the detections of 89% and a specificity of 85%. However, this model was developed and evaluated specifically for lameness detection and only severe lameness cases were included in the analysis. Moreover, milk yield data, used as input variables, were more correlated with lameness than rumination and neck activity data. Thus, the good results of this model may only partly be explained by the use of rumination and activity as input variables, and mainly by the inclusion of milk yield related data, which were not included in our detection algorithms.

Several method developments are original in this study. Unlike many studies focused on the detection of one disease, a generic approach is proposed to evaluate broad spectrum monitoring tools, e.g. monitoring tools likely to detect a large variety of HD. It is based on detection status determination, in time-windows around HD detection for sensitivity calculation, and in control time-windows randomly sampled away from other control time-windows and from HD detection for specificity calculation. This random sampling process is a way to limit autocorrelation in detector data and to address the class imbalance problem classically met when working on detection algorithms (Galar et al., 2012), and due in this study to the low frequency of HD when compared to healthy periods. Our results also underline the interest of analyzing separately HD occurring around calving and HD occurring during lactation or dry period. Indeed, the specificity of CUSUM detections was sharply lower for HD_{pp} compared with HD, probably due to the physiological drops in RT often met around calving (Schirrmann et al., 2013). As sensor data are often not available during the dry period, and as physiological modifications occur at calving (Adams et al., 2013), the informative value of the detections around calving are often not fully evaluated in other studies. However a high proportion of HD occur around calving (Gernand et al., 2012). Thus, it is not ideal to exclude the peripartum period from the analyses, because HD such as milk fever or acute metritis, mainly met just after calving, should be detected for appropriate treatment. The principle of most studies dealing with detection model development is to look for the best trade-off between sensitivity and specificity of the detections. In our study, due to the low frequency of HD, reaching a high specificity to limit FP detections and to enable the practical use of the tool by farmers was more important than reaching a high sensitivity.

Thus, another original approach in this study was to fix either the specificity or the sensitivity level instead of researching the best trade-off between sensitivity and specificity. The false detection rate, the PPV and the overall sensitivity were then examined to choose the best settings.

The PPV obtained were below 10% in our sample. Firstly, this may be due to a lack of specificity of the detections. Many events, such as a modification of management routine (feeding, milking, changing pen), may have influenced the behavior of the cows and provoked FP detections. For example at drying off during winter, newly dried cows were transferred into a smaller pen in which the amount of corn silage offered was reduced. This might have provoked sudden drops in RT or in NA. All the changes in individual cows' management routines could not be recorded by farm staff daily over the study. Thus some detections linked to changes in management routines may have been counted as FP events. Moreover, main factors affecting RT include diet digestibility, NDF intake (Adin et al., 2009), forage quality (Welch and Smith, 1970), heat stress (Soriani et al., 2013b) and regrouping (Schirmann et al., 2011). Thus, some FP detections may have been due to modifications in those factors that were not completely controlled over the study. Part of FP detections might also have been due to real HD provoking drops in RT or NA but not recorded by farm staff as the cow coped with this corresponding HD, which remained subclinical. Another source of FP detections might be the lack of sensitivity of farm staff detections. Indeed, Timsit et al. (2011b) found that only 26% of the bulls presenting hyperthermia after arrival at fattening operations were detected as diseased by farm staff through visual observation. Similarly, Leach et al. (2010) found that, on average, there was a 30% underestimation of lameness prevalence by farm staff compared with standard notation by researchers on a 222 dairy farm sample in the UK. Thus, the reference method used to classify an animal as diseased or healthy is far from being ideal. In particular, it is not possible to exclude that in our study, some events were misclassified into FP instead of TP due to the imperfection of HD detection by farm staff. A second reason for low PPV was the low frequency of cow days where it was relevant to detect HD. Thus, due to the low frequency of HD, PPV remains low if very high sensitivity and specificity levels are not reached. For mastitis detection by AMS, the International Standard ISO/FDIS 20966 (ISO, 2007) recommends a sensitivity of 80% combined with a specificity better than 99%. These recommendations are not often reached when short time-windows are used for detection validation (Hogeveen et al., 2010, Rutten et al., 2013). Moreover, even with these levels of sensitivity and specificity, the PPV obtained may be lower than 30% (Kamphuis et al., 2008). In our study, we chose to test a range of proportions of diseased cow days (0.5 to 5%). This range of diseased cow days was explored to take into

account the variability of health status across herds and the potential under detection of HD by farmers. Indeed, the prevalence of these cow days in a herd may vary according to the overall health status of the herd but also according to the level of grouping together of calvings, as a majority of HD occurs during early lactation (Gernand et al., 2012). For some HD such as bovine respiratory disease (Timsit et al., 2011b) or lameness (Leach et al., 2010), farmers have been shown to detect only a minor part of diseased animals. Thus the true prevalence of HD in a herd is probably higher than the prevalence calculated from farmers' detections. Moreover, comparing values of sensitivity, specificity, PPV and NPV obtained in different studies must be done carefully because the way of counting TP, FP, TN, FN and calculating PPV and NPV is not standardized between studies. Finally, even if the PPV of a detection was very low, it should be noted that in practice, farmers integrate information from information systems, such as milk yield, visual observation, reproduction and health management tools, before deciding to make a clinical examination of a cow that has triggered a detection. Thus for health monitoring tools, the decision of examining an animal will rarely be based exclusively on a detection by the monitoring tool.

An effort was made throughout our work to discuss its practical implications for the users of the detection algorithms used. Indeed, we based our analysis on user transparent measures of the quality of the detections. The PPV is important for the farmers as it indicates the proportion of detections that would correspond to diseased animals. The false detection rate is also a simple indicator linked to the specificity of the detections. Due to methodological difficulties or maybe to their often unfavorable values, these indicators are rarely reported in health monitoring device studies. Another result of this study is that the type of detection, i.e. drop of RT or of NA, will not be an indication of the type of HD the farmer should look for as a priority during the clinical examination of an animal detected by the algorithms. Indeed, there is not a systematically better detection of one type of HD by the CUSUM on RT or by the CUSUM on NA. Depending on the severity of the HD, the sensitivity of the CUSUM on RT may be higher than the sensitivity of the CUSUM on NA or vice-versa. For instance, with CUSUM setting 2, more impaired general condition episodes without hyperthermia were detected by the CUSUM on RT (Se = 53.8% vs. 43.6%), whereas more impaired general condition episodes with hyperthermia were detected by the CUSUM on NA (Se = 62.5% vs. 45.8%). Over the study, 25% of targeted HD were excluded from the analysis due to excessive missing sensor values. Thus, the sensitivity of the detections was overestimated in our study, as the detection of HD occurring during dry period or during periods with excessive missing sensor data is desirable from a farmer's point of view. A distinction between HD occurring during the dry period, where the device did not record data,

and HD occurring during periods with excessive missing data due to technical problems on the device, could not be established in our study. However, the last version of the commercial device uses wifi technology for transmitting data from the RT and NA logger to the computer that analyses data. Thus commercial HD detections are now available during the dry period if dry cows are located close to the wifi antenna (in the same barn or on pasture immediately outside the barn).

For the more frequent HD, the high proportion of events detected by the CUSUM just around the detection by farm staff (days -1, 0 and 1) seems to indicate that CUSUM detections observed around HD detections by farm staff were actually linked to the occurrence of these HD. Our results indicate that approximately 50% of the detected events would be detected before the detection by farm staff, with 5 to 15% of HD detected 3 d before. Similarly, Timsit et al. (2011a) found that ruminal hyperthermia episodes could be detected 12 to 90 h before the onset of clinical signs of respiratory disease in a bull fattening unit. For these HD detected earlier, a better cure rate might be obtained after treatment, compared with the situation where no monitoring tool is used. These early detections would contribute to improve animal welfare and to limit the economic impact of HD.

8. Conclusions

The present study shows that the monitoring of drops in RT and in NA could be useful for the detection of a large variety of HD. For an acceptable level of specificity (90.1%), on average 39.6% of the HD observed by farm staff were detected with 34.9% of mastitis, 30.2% of lameness, 63.5% of impaired general condition episodes without identified cause detected respectively. However, most of the HD of high severity were detected with a sensitivity higher than 55%. In spite of this, these automated detections would only be a complement to visual observation for HD detection. Indeed, due to the moderate specificity of the algorithms (90%) and to the low prevalence of HD, the PPV of a detection of a drop in RT or in NA remained below 10% in our study sample. Thus the overall informative value of the detection of a drop in RT or in NA was not sufficient to be as efficient as visual observation for HD detection. However, in this study we used farmer detection as a reference method. But farmers have been shown to under detect some HD and we can assume that the sensitivity of HD detection varies over time according to the time and attention the farmer can pay to his cows. Future work will

focus on the development of an appropriate reference method to evaluate such a monitoring tool, likely to detect a broad spectrum of health events.

9. Acknowledgments

We gratefully acknowledge the contributions made by farm staff (“la station expérimentale laitière de la chambre d’agriculture du Maine et Loire (Les Trinottières)” and “la station expérimentale laitière des chambres d’agriculture de Bretagne (Trevarez)”) in gathering the data used in this research. “Chambre d’Agriculture du Maine et Loire”, “chambre d’abriculture de Bretagne” and “Institut de l’élevage” are also acknowledged for their financial and technical support. We are also grateful to Milkline srl (Podenzano, Piacenza, Italy) for financial and technical support. We would like to thank Aurélien Madouasse (Oniris-UMR Bioepar, Nantes, France) and Francois Guillaume (Evolution, Rennes, France) for their critical comments on the manuscript. This research is supported by Evolution (Rennes, France).

References

- Adams, A. E., F. J. Olea-Popelka, and I. N. Roman-Muniz. 2013. Using temperature-sensing reticular boluses to aid in the detection of production diseases in dairy cows. *J Dairy Sci* 96:1549-1555.
- Adin, G., R. Solomon, M. Nikbachat, A. Zenou, E. Yosef, A. Brosh, A. Shabtay, S. J. Mabweesh, I. Halachmi, and J. Miron. 2009. Effect of feeding cows in early lactation with diets differing in roughage-neutral detergent fiber content on intake behavior, rumination, and milk production. *J Dairy Sci* 92:3364-3373.
- Bewley, J. M. 2009. Precision Dairy Farming: the next dairy marvel ? in 60th Annual Meeting of the European Association for Animal Production. Vol. Abstract 5652, Barcelona, Spain.
- Chapinal, N., A. M. de Passille, J. Rushen, and S. Wagner. 2010. Automated methods for detecting lameness and measuring analgesia in dairy cattle. *J Dairy Sci* 93:2007-2013.
- de Koning, C. J. A. M. 2010. Automatic milking-Common practice on dairy farms. Pages 52-67 in Proc. First North. Am. Conf. Precision Dairy Management, Toronto, Canada.
- EFSA (European Food Safety Authority). 2009. Scientific Opinion of the Panel on Animal Health and Welfare on a request from European Commission on welfare of dairy cows. Effects of farming systems on dairy cow welfare and disease. in *The EFSA Journal*. 1143. 1-38.
- Fitzpatrick, C. E., N. Chapinal, C. S. Petersson-Wolfe, T. J. DeVries, D. F. Kelton, T. F. Duffield, and K. E. Leslie. 2013. The effect of meloxicam on pain sensitivity, rumination time, and clinical signs in dairy cows with endotoxin-induced clinical mastitis. *J Dairy Sci* 96:2847-2856.
- Fourichon, C., H. Seegers, N. Bareille, and F. Beaudeau. 1999. Effects of disease on milk production in the dairy cow: a review. *Prev Vet Med* 41:1-35.
- Galar, M., Ferna, x, A. ndez, E. Barrenechea, H. Bustince, and F. Herrera. 2012. A Review on ensembles for the class imbalance problem: bagging-, boosting-, and hybrid-based approaches. *Systems, Man, and Cybernetics, Part C: Applications and Reviews, IEEE Transactions* 42:463-484.
- Gernand, E., P. Rehbein, U. U. von Borstel, and S. König. 2012. Incidences of and genetic parameters for mastitis, claw disorders, and common health traits recorded in dairy cattle contract herds. *J. Dairy Sci.* 95:2144-2156.

Hogeveen, H., C. Kamphuis, W. Steeneveld, and H. Mollenhorst. 2010. Sensors and clinical mastitis--the quest for the perfect alert. *Sensors (Basel)* 10:7991-8009.

Holman, A., J. Thompson, J. E. Routly, J. Cameron, D. N. Jones, D. Grove-White, R. F. Smith, and H. Dobson. 2011. Comparison of oestrus detection methods in dairy cattle. *Vet Rec* 169:47.

ISO (International Organization for Standardization). 2007. Automatic milking installations—Requirements and testing. Annex C: Example of methods of evaluating detection systems for milk deemed as abnormal due to blood or to changes in homogeneity. ISO 20966:2007. Geneva, Switzerland.

Jensen, M. B. 2012. Behaviour around the time of calving in dairy cows. *Appl. Anim. Behav. Sci.* 139:195-202.

Kamphuis, C., R. Sherlock, J. Jago, G. Mein, and H. Hogeveen. 2008. Automatic detection of clinical mastitis is improved by in-line monitoring of somatic cell count. *J Dairy Sci* 91:4560-4570.

Kamphuis, C., B. DelaRue, C. R. Burke, and J. Jago. 2012. Field evaluation of 2 collar-mounted activity meters for detecting cows in estrus on a large pasture-grazed dairy farm. *J Dairy Sci* 95:3045-3056.

Leach, K. A., H. R. Whay, C. M. Maggs, Z. E. Barker, E. S. Paul, A. K. Bell, and D. C. Main. 2010. Working towards a reduction in cattle lameness: 1. Understanding barriers to lameness control on dairy farms. *Res Vet Sci* 89:311-317.

Luo, Y., Z. Li, and Z. Wang. 2009. Adaptive CUSUM control chart with variable sampling intervals. *Comput. Stat. Data Anal.* 53:2693-2701.

Radostits, O. M., C. C. Gay, K. W. Hinchcliff, and P. D. Constable. 2007. *Veterinary Medicine: A textbook of the diseases of cattle, horses, sheep, pigs, and goats*. 10th ed. Saunders Elsevier, New York

Reith, S. and S. Hoy. 2012. Relationship between daily rumination time and estrus of dairy cows. *J Dairy Sci* 95:6416-6420.

Rutten, C. J., A. G. Velthuis, W. Steeneveld, and H. Hogeveen. 2013. Invited review: sensors to support health management on dairy farms. *J Dairy Sci* 96:1928-1952.

Schirmann, K., M. A. von Keyserlingk, D. M. Weary, D. M. Veira, and W. Heuwieser. 2009. Technical note: Validation of a system for monitoring rumination in dairy cows. *J Dairy Sci* 92:6052-6055.

Schirmann, K., N. Chapinal, D. M. Weary, W. Heuwieser, and M. A. von Keyserlingk. 2011. Short-term effects of regrouping on behavior of prepartum dairy cows. *J Dairy Sci* 94:2312-2319.

Schirmann, K., N. Chapinal, D. M. Weary, L. Vickers, and M. A. von Keyserlingk. 2013. Short communication: Rumination and feeding behavior before and after calving in dairy cows. *J Dairy Sci*.

Seegers, H., C. Fourichon, and F. Beaudeau. 2003. Production effects related to mastitis and mastitis economics in dairy cattle herds. *Vet Res* 34:475-491.

Soriani, N., E. Trevisi, and L. Calamari. 2012. Relationships between rumination time, metabolic conditions, and health status in dairy cows during the transition period. *J Anim Sci* 90:4544-4554.

Soriani, N., G. Panella, and L. Calamari. 2013a. Rumination time during the summer season and its relationships with metabolic conditions and milk production. *J Dairy Sci* 96:5082-5094.

Soriani, N., G. Panella, and L. Calamari. 2013b. Rumination time during the summer season and its relationships with metabolic conditions and milk production. *J Dairy Sci* 96:5082-5094.

Timsit, E., S. Assie, R. Quiniou, H. Seegers, and N. Bareille. 2011a. Early detection of bovine respiratory disease in young bulls using reticulo-rumen temperature boluses. *Vet J* 190:136-142.

Timsit, E., N. Bareille, H. Seegers, A. Lehebel, and S. Assie. 2011b. Visually undetected fever episodes in newly received beef bulls at a fattening operation: occurrence, duration, and impact on performance. *J Anim Sci* 89:4272-4280.

Van Hertem, T., E. Maltz, A. Antler, C. E. Romanini, S. Viazzi, C. Bahr, A. Schlageter-Tello, C. Lokhorst, D. Berckmans, and I. Halachmi. 2013. Lameness detection based on multivariate continuous sensing of milk yield, rumination, and neck activity. *J Dairy Sci* 96:4286-4298.

Welch, J. G. and A. M. Smith. 1970. Forage quality and rumination time in cattle. *J Dairy Sci* 53:797-800.

Partie II. Intérêt du monitoring des baisses de durées de rumination et d'activité générale pour la détection des chutes de production laitière chez les vaches laitières

1. Résumé de la Partie II

L'évaluation des performances de systèmes de détection des troubles de santé basés sur les données issues de capteurs est importante pour leur utilisation optimale en ferme. Cependant, les signes cliniques de certaines maladies sont parfois frustes. En conséquence, les éleveurs détectent mal certains troubles de santé. Cela peut limiter l'utilisation de la détection des troubles de santé par observation visuelle de l'éleveur, pour déterminer avec précision l'état de santé des animaux (sains ou malades) lors de l'évaluation d'un système de monitoring de la santé. C'est pourquoi il pourrait sembler intéressant d'utiliser des signes indirects de maladie, par exemple des baisses de production laitière, pour évaluer un modèle de détection. Dans cette étude, les performances d'algorithmes de type 'cumulative sum' appliqués sur des données d'activité générale et de durées de rumination ont été évaluées, pour la détection (i) de troubles de santé détectés par des éleveurs et/ou (ii) de chutes importantes de production laitière.

Au total, 171 vaches Holstein de 2 fermes expérimentales de l'ouest de la France ont été suivies pendant deux ans et demi. Les troubles de santé étaient enregistrés par les éleveurs de manière standardisée. Les chutes importantes de production laitière ont été détectées rétrospectivement à l'aide d'un test 'cumulative sum'. Les baisses d'activité générale ou de rumination, mesurées à l'aide d'un capteur situé au cou de l'animal, ont été détectées en utilisant le même type d'algorithmes. Au cours de l'étude, 501 troubles de santé ont été détectés par les éleveurs, parmi lesquels 94 étaient concomitants à une baisse importante de production laitière et 407 n'étaient pas associés à une baisse importante de production laitière. De même, 263 chutes importantes de production laitière n'étaient pas associées à la détection d'un trouble de santé par les éleveurs. Respectivement 18,4, 29,2 et 56,5% des troubles de santé non concomitants aux baisses importantes de production laitière, des baisses importantes de production laitière non concomitantes et des baisses de production laitière concomitantes aux troubles de santé ont été détectés par le test 'cumulative sum' sur l'activité ou la rumination, avec un niveau élevé de spécificité (95,2%). Parmi les 103 chutes importantes de production laitière détectées par le test 'cumulative sum' sur l'activité ou la rumination, 60% ont été détectées le jour du début de la baisse de production laitière ou dans les 3 jours précédents.

Nos résultats montrent que de nombreuses chutes importantes de production laitière (0,7 par lactation en moyenne) n'étaient pas associées à la détection d'un trouble de santé par les éleveurs. Ce résultat était attendu en raison du manque de sensibilité des éleveurs pour la détection de certaines maladies. Vingt-neuf pourcent de ces chutes importantes de production laitière étaient associées à une alarme déclenchée par le test 'cumulative sum' sur l'activité ou la rumination. Lorsqu'elles sont accessibles, les baisses de production laitière devraient être prises en compte lors de l'évaluation d'un système de monitoring des troubles de santé, en particulier pour évaluer correctement la spécificité des alertes de santé fournies par le système.

Interest of monitoring drops in rumination time or in neck activity for the detection of milk production losses in dairy cows

En préparation.

2. Abstract

The evaluation of the performance of health disorders detection models based on sensor data is a major concern for an optimal use in dairy farms. However, clinical signs of disease are sometimes unapparent or slight. Thus, farmers have shown lack of sensitivity for the detection of particular health disorders (HD), which might compromise the use of visual appraisal by farm staff to correctly classify animals as healthy or sick when evaluating a monitoring device. Therefore, using indirect manifestation of disease such as drop in milk production for evaluating a detection model could be a seducing alternative. In this study, the ability of cumulative sum algorithms applied on neck activity (NA) and rumination time (RT) data to detect (i) HD recorded by farm staff and/or (ii) major drops in milk production was evaluated. Overall, 171 dairy cows from 2 research herds in Western France were enrolled in a two and half year study. HD were recorded by farm staff on a standard form. Major drops in milk production were detected using a cumulative sum test. Drops in NA or in RT measured by a neck sensor were detected using similar algorithms. Over the study period, 501 HD were detected by farm staff, among which 94 occurred around a major drop in milk production and 407 were not associated with a major drop in milk yield. Moreover, 263 major drops in milk production were not associated with a HD detection by farm staff. Overall, 18.4, 29.2 and 56.5% of independent HD,

independent drops in milk production and concomitant HD and drop in milk production were detected by the cumulative sum test on NA or RT respectively, for a high specificity (95.2%). Among the 103 major drops in milk production detected by the cumulative sum on RT or NA, 60% were detected on the day of the beginning of the drop in milk production or within the previous 3 days. Our results show that many major drops in milk production (0.7/lactation) were not associated with the detection of a HD by visual appraisal by farm staff. This result is consistent with a potential lack of sensitivity of visual appraisal by farm staff for HD detection. Twenty-nine percent of these drops in milk production were associated with an alarm triggered by the cumulative sum test on RT or NA. When available, drops in milk production should be considered when evaluating an HD detection model, especially to correctly estimate the specificity of the model.

3. Introduction

Precision dairy farming is anticipated as a mean of increasing the profitability of dairy industry. Bewley (2010) defined precision farming as “the use of technologies to measure physiological, behavioral, and production indicators on individual animals to improve management strategies and farm performance”. Monitoring tools recently developed to assist dairy farmers in the management of the reproduction and the health of their cows clearly fit this definition. These tools enable farmers to manage larger herds with lower labor requirements (de Koning, 2010), as instead of managing each cow or group of cows, the management is mainly focused on individual cows deviating from the group or from their usual values (Schulze et al., 2007). However, the usefulness of these tools strongly depends on their technical performance. Indeed, a lack of sensitivity or specificity of the detections may alter farmers' confidence in the tool and may lead to its underuse. Farm profitability, animal welfare and optimal use of antibiotic treatments may also be negatively impacted if an insufficiently sensitive and specific tool is used to detect diseased animals.

The evaluation of the performance of a continuous monitoring tool requires determining the health status of animals over the time using a reference method. It is classically performed by practicing specific tests to detect particular HD such as mastitis, lameness or metabolic disorder. For instance, somatic cell count is often used as a reference method for mastitis detection (Rutten et al., 2013). On the other hand, health status may also be determined indirectly as diseases may impact animal behavior and provoke decreased DMI or milk production for instance (Bareille et

al., 2003). As a consequence, several studies used milk yield, water intake or DMI to detect disease onset in dairy cows (Huybrechts et al., 2014, Lukas et al., 2008, Lukas et al., 2009, Miekley et al., 2012). In particular, milk losses have been shown to be associated with a large variety of health disorders (Fourichon et al., 1999).

Recently, a rumination and activity sensor (HR-Tag, SCR Engineers Ltd., Netanya, Israel) has been developed. This sensor is supported by a neck collar and continuously records rumination time (**RT**) and neck activity (**NA**) of each cow of the herd. Sensor data are used to generate lists of cows that need special attention concerning estrus (in case of overactivity) or health (drops in activity or rumination time, or both). The spectra of HD potentially detectable with this device is largely unknown but is expected to be broad as sensors' data, i.e. RT and NA, may be affected by a large variety of health disorders (Radostits et al., 2007). As a consequence, correctly evaluating its performance raises several issues. An evaluation of the tool using a specific test for each targeted health disorder, i.e. a test for mastitis, a test for lameness, a test for metritis, a test for ketosis, would be very expensive and time consuming. Using visual appraisal by farm staff as a reference method is an appealing idea. However farmers have been shown to greatly lack sensitivity for the detection of HD such as pneumonia (Timsit et al., 2011) or lameness (Fabian et al., 2014). Indeed, Timsit et al. (2011) found that after arrival of bulls at 3 fattening operations, only 26% of fever episodes were detected by feedlot personnel. Similarly, Fabian et al. (2014) found that only 27% of cows with reduced mobility due to lameness were identified as lame by farm staff. Thus using visual observation by farm staff to classify animals as healthy or diseased may lead to a lot of misclassifications, and particularly to falsely consider some detections as false positives (in case of alarm triggered by the monitoring tool, corresponding to a "true" health disorder, but without concomitant detection by farm staff). We assumed that using an indirect indicator of disease such as milk yield, which is relatively cheap and easy to collect daily, could be relevant to evaluate the performance of this monitoring device.

The objective of this study was to evaluate the ability of a one-sided CUSUM applied on RT and NA measured by HR-Tag, in detecting major drop in milk production (**DMP**).

4. Materials and methods

Cows and Herd Management

The data used in this study were collected on 2 research farms in Western France where management practices were representative of classical farming conditions of this area. Cows were milked twice daily and fed a corn silage based TMR or a pasture based diet, depending on the period of the year. On farm 1, (Les Trinottières: Maine-et-Loire, France) approximately 115 Holstein cows were involved between September 01, 2010 and August 31, 2013. Cows were managed in 2 batches: 1 (60 cows) was fed a TMR containing 75% corn silage and 25% concentrate (zero-grazing) and the other batch was fed the same TMR from November to April and a diet containing 50% grass and 50% corn silage from April to November. On farm 2, (Trevarez: Brittany, France), 56 Holstein cows were involved between November 04, 2009 and January 11, 2012. Cows were fed a corn silage based TMR during winter or a pasture based diet from spring to autumn. Cows were housed in cubicles during winter on both farms, and average milk yield was 30.6 kg/d (farm 1) and 25.9 kg/d (farm 2).

Health Disorder Recording

Throughout the study period, all cows were monitored daily by farm staff for signs of disease according to standardized protocols. Farm staff had no access to any automatic detection system during the study. Thus, RT or NA could not be used to identify cows with HD. Once detected, the nature, severity and date of occurrence of HD were recorded on a standardized form (table 3-II-1). On each farm, one person was in charge of recording HD and the author ensured homogenous recording between farms.

Table 3-II-1: standard form used by farm staff for recording health disorders. Definition of the type and severity of the health disorders

	Severity	Description
Mastitis	1	Clots in milk only
	2	Clots in milk, udder inflammation, red and painful quarter
	3	Mastitis with general clinical signs: hyperthermia or drop in milk production or decreased appetite
Lameness	1	Preventive hoof treatment
	2	The cow is obviously lame on 1 or more legs. An observer will in most cases be able to tell which leg is affected
	3	The cow is obviously lame on 1 or more legs. The cow is unable, unwilling, or very reluctant to bear weight on the affected leg
Milk fever	1	Treatment of a cow suspected to begin a milk fever
	2	Treatment of a cow down with milk fever symptoms
Retained placenta		Retained placenta more than 24 h post-partum with general clinical signs: hyperthermia or drop in milk production or decreased appetite
Metritis	1	Vulvar muco-purulent material discharge before 21 d postpartum
	2	Endometritis: vulvar discharge with purulent material after 21 d postpartum, without alteration of general condition
	3	Acute metritis: vulvar discharge with purulent material before 21 d postpartum with impaired general condition: hyperthermia or drop in milk production or decreased appetite
Impaired general condition	1	Impaired general condition without identified cause: drop in milk production or decreased appetite without hyperthermia
	2	Impaired general condition without identified cause: drop in milk production or decreased appetite with hyperthermia
Other health disorder		Description as precise as possible of the health disorder

Detection of Drops in Milk production

The objective was to detect major DMP, potentially associated with health disorders. Individual yield was recorded daily throughout the study using milkmeters.

In the first step, expected milk yield was predicted daily using a method based on locally weighted polynomial regression (proc LOESS, SAS version 9.2; SAS Institute Inc., Cary, NC), a smoothing method developed by Cleveland and Grosse (1991). At each point in the data set, a polynomial is fit to a subset of the data within a chosen neighborhood of the point whose response is being estimated. In this study, a second order polynomial was fit before 20 DIM to

ensure a better fit to data during the first part of the lactation, and a first order polynomial was fit after 20 DIM. The fraction of data which is used for the regression in each local neighborhood is called the “smoothing parameter”, which determines the smoothness of the estimated curves. The smaller the smoothing parameter is, the closer the line fits the original data points. The smoothing parameter is usually estimated based on AIC criteria. In addition, in the LOESS method, the individual time distances of the observations are taken into account. More weight is given to points temporally close to the point whose response is being predicted and less weight to points further away. In this study, the proc LOESS was used to estimate the expected milk yield in absence of perturbation. This was done by over smoothing lactation curves. An optimal smoothing parameter was estimated for each lactation, based on AIC criteria. Then, a smoothing parameter equal to 0.2, i.e. 10 times the median of smoothing parameters obtained in step 1 for all the lactations included in the analysis, was applied for all the lactations. In order to correctly predict the expected milk yield during DMP, local reweighting was performed. During this operation, several smoothing iterations are performed, and the weight given to an observation decreases as the residual of the fitted value at the previous iteration increases (Cleveland and Grosse, 1991). In this study, the number of iterations was set to 5.

In the second step, after prediction of expected milk yield, a one-sided cumulative sum (CUSUM) test (Luo et al., 2009) was applied retrospectively for each lactation on daily residuals (observed – predicted yield) to detect DMP. The CUSUM test accumulates the difference (error) between the predicted and the observed milk yield when this difference is positive. The CUSUM used in this study was computed for each cow as:

$$\begin{cases} C_0 = 0 \\ C_t = \max(0 ; C_{t-1} + \varepsilon_t - k) \end{cases}$$

Where C_t is the CUSUM value at time ‘t’, ε_t is the difference between the predicted and the observed values at time ‘t’, k is equal to $\sigma_{(\text{pred-obs})}$, where $\sigma_{(\text{pred-obs})}$ is the standard deviation of the error (difference between observed and predicted values) calculated for each cow for the considered lactation. Decision interval (h) was set to $3.5 \times \sigma_{(\text{pred-obs})}$. These two values were computed as functions of $\sigma_{(\text{pred-obs})}$ to take into account the variability of the standard deviation of residuals across cows. An alarm was triggered when the CUSUM value exceeded a given decision interval (h). SAS 9.2 (SAS Institute Inc., Cary, NC) was used for algorithm development and data analyses.

For each CUSUM alarm a DMP beginning the first day when residual was negative before CUSUM alarm and ending the last day when residual was negative after CUSUM alarm was defined. Drops lasting for less than 3 days were deleted to avoid including outliers in the analysis. The residuals during the drop were used to calculate total relative loss = total loss (kg)/mean predicted yield (kg) and maximum relative loss = maximum daily loss (kg) / $\sigma_{(\text{pred-obs})}$ (kg) (figure 3-II-1). Then, only major drops were selected, based on their amplitude and total loss. Only losses with either a maximum relative loss higher than the median of all detected maximum relative losses (3.7) or with total relative loss higher than the median of all detected losses (0.82) were considered as major drops and thus relevant to detect. For retained losses, total loss (kg), mean daily loss (kg/d), maximum daily loss (kg/d) and loss duration (d) were calculated from residuals (figure 3-II-1).

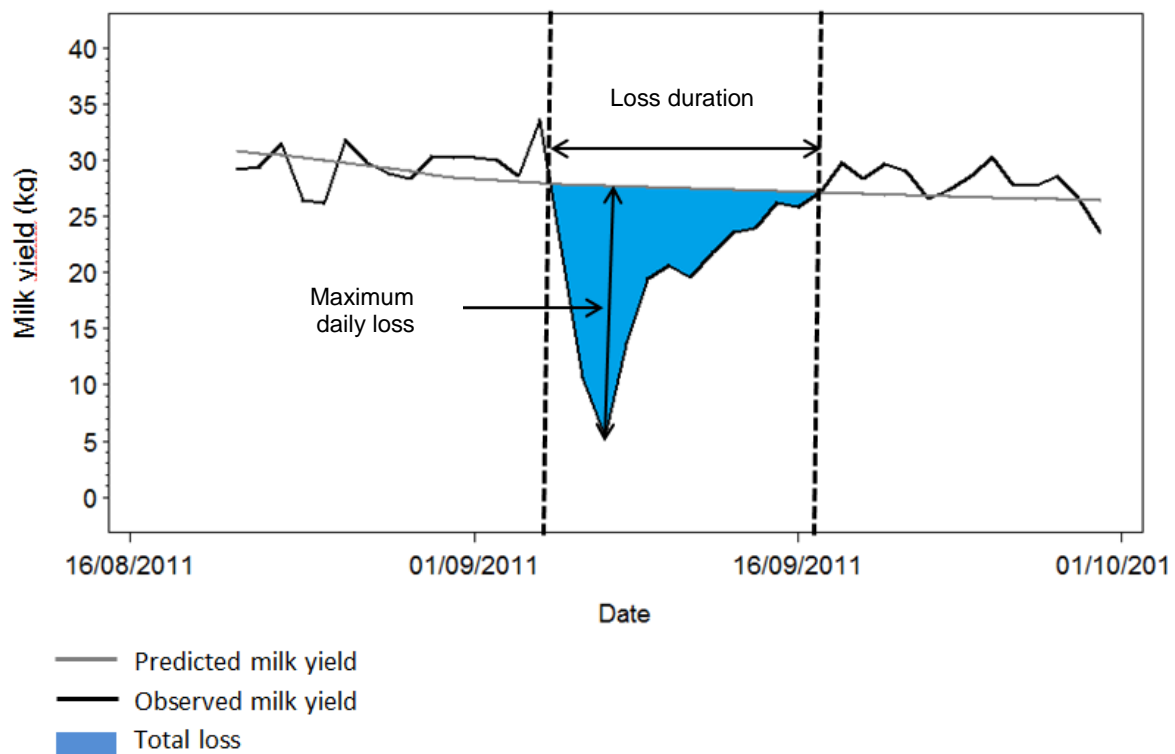


Figure 3-II-1: example of one loss characteristic calculation.

Concomitancy between Farm Staff Detections and Drops in Milk

In order to describe the concomitancy between HD detection by farm staff and major DMP detected by the CUSUM test, 10 d time-windows centered on the detection by farm staff were

computed. The presence or absence of major DMP in these time-windows were counted to determine whether DMP were concomitant or not with HD respectively. Similarly, the number of major DMP outside these time-windows was counted. Only HD and DMP occurring after 5 DIM and during lactation were included in the analysis. Preventive hoof treatments (n = 63) were ignored for the analysis and DMP occurring in a 10 d time-window around a preventive hoof treatment were not considered as concomitant.

Sensors' Data

At calving or a few days before, depending on standard procedures in the farm, cows were fitted with a neck collar supporting a NA and RT logger (HR-Tag, SCR Engineers Ltd., Netanya, Israel). The logger contains a microphone that records RT continuously and summarizes it into 2 h blocks (Schirmann et al., 2009). It also contains an accelerometer that, after algorithmic treatment, summarizes NA into 2 h blocks (Kamphuis et al., 2012). RT and NA data were stored in the logger and transmitted to a computer (DataFlow software, SCR Engineers Ltd., Netanya, Israel) when cows walked below an infrared antenna located in the milking parlor or over a trough. Thus, new data were typically available twice a day after milking and data were often not available during dry period.

Algorithms for the Detection of Drops in RT and NA

As the commercial algorithms developed by the manufacturer for HD detection with the device used in this study were confidential, the authors of the present study developed their own algorithms. The algorithms were retrospectively applied on sensor data.

A method based on a one-sided cumulative sum (CUSUM) test was used to detect drops in RT or NA. The algorithm developed is described in details elsewhere (chapter 3.I). Briefly, firstly reference RT and NA were calculated for each cow and for each 2 h block of the day (12 blocks/d) as the mean RT and NA respectively, recorded for this cow at the same 2 h time block of the day during the past 7 d. Secondly, an alarm was triggered when the CUSUM value

exceeded a given decision interval (h). The CUSUM used in this study was computed for each cow as:

$$\begin{cases} C_0 = 0 \\ C_t = \max(0 ; C_{t-1} + \varepsilon_t - k) \end{cases}$$

Where C_t is the CUSUM value at time t , ε_t is the difference between the reference RT or NA values and the observed values at time t , k is equal to $b \times \sigma_{(\text{ref-obs})7d}$, where b is a constant and $\sigma_{(\text{ref-obs})7d}$ is the standard deviation of the error (difference between observed and reference values) calculated for RT or NA for each cow on all the values of the last 7 d, irrespective of the time block of the day. This last parameter was added to take into account the high variability of RT and NA between cows. b value was set to 0.7, as this value gave the highest level of specificity for a given level of sensitivity, and gave the highest level of sensitivity for a given level of specificity in a previous study (Clément et al., submitted) using farm staff detection as a reference method. Decision intervals (h) enabling to reach a high specificity in this preliminary experiment, i.e. 100 for the CUSUM on RT and 75 for the CUSUM on NA, were chosen.

Informative Value of Sensors' Detections

As a monitoring tool may detect a HD a few hours or a few days before or after the detection by the reference method, the definition of a time-window in which a detection by the tool is likely to be caused by the event to detect is necessary for informative value calculation (Hogeveen et al., 2010). Then, in our study, an 8 d time-window starting 72 h (3 d) before the beginning of each major DMP and ending 120 h (5 d) after was defined for sensitivity calculation. This 8 d time-window was called the validation time-window. A validation time-window was excluded from the analysis if more than 33% of RT or NA data were missing in this time-window. A validation time-window was considered as a True Positive (**TP**) if an alarm was triggered by the CUSUM on RT or on NA during this time-window. A validation time-window was considered as a False Negative (**FN**) if no alarm was triggered. The sensitivity of the CUSUM detections was calculated as $Se = TP/(TP + FN)$. Mean total loss, mean daily loss, maximum daily loss and duration of the losses detected by the CUSUM on RT or NA and of the losses not detected by the CUSUM were compared using a t-test (SAS 9.2). If a HD was detected by farm staff during a

validation time-window, the HD and the DMP were counted as concomitant. Otherwise, the DMP was counted as independent.

For specificity calculation, three 8 d control time-windows were randomly sampled (proc surveysselect, SAS 9.2) for each cow with sufficient sensor data (more than 3.5 months of data). To be validated, these 8 d control time-windows had to start after 16 DIM, to start at least 20 d before or 20 d after a DMP or a HD detection by farm staff, to be distant from another control time-window by a minimum of 20 d, and to contain less than 33% of missing RT or NA data. A control time-window was considered as a False Positive (**FP**) if an alarm was triggered by the CUSUM on RT or on NA during this time-window. A control time-window was considered as a True Negative (**TN**) if no alarm was triggered during this time-window. The specificity was calculated as $Sp = TN / (TN + FP)$. Among the 222 cows of the study with more than 3.5 months of data, 202 could be correctly sampled due to the constraints imposed to validate the control time-windows. Thus, specificity calculations were made on 606 (3 x 202) 8 d control time-windows.

For sensitivity and specificity calculations, an exception was made for the detections occurring for the CUSUM on RT data between 8 h before and 26 h (estrus time-window) after the beginning of an overactivity episode detected by the commercial algorithm for estrus detection associated with the device (Kamphuis et al., 2012). As a decrease in RT is often observed around estrus (Reith et Hoy, 2012), these drops in RT detected in the estrus time-window around activity peaks (overactivity algorithm value > 5.2) were neither counted as TP (if in a validation time-window) nor as a FP (if in a control time-window) but instead were ignored.

For HD not concomitant with a major DMP, the presence or absence of alarms triggered by the CUSUM on RT or NA in 8 d (-3 d ; +5 d) time-windows around HD detected by farm staff was also computed and the sensitivity of the alarms was calculated as explained above.

Time-lag between Sensors' Detections and Drops in Milk

Another important aspect when evaluating a monitoring tool is the time lag between the alarms triggered by the tool and the detections by the reference method. Indeed, an early detection may lead to earlier treatment with improved cure rates. To investigate this issue, the time-lag between

the first detection by the CUSUM algorithm on RT or on NA data in a given validation time-window and the beginning of the DMP was calculated for each detection of a major DMP.

5. Results

Concomitancy between Farm Staff Detections and Drops in Milk

Overall, 347 major DMP were detected during the study period. The mean number of drops per cow was 1.9 (range 1-5) and the mean number of drops per lactation was 0.9. Descriptive statistics concerning the losses included in the analysis are displayed in table 3-II-2. Mean values for total losses, maximum daily losses, mean daily losses and loss durations were respectively 40.7 kg, 9.8 kg/d, 4.0 kg/d and 10.7 d.

Table 3-II-2: descriptive statistics of major drops in milk production detected in the study (n = 347)

	Mean	SD	Min	Max
Total loss (kg)	40.7	30.05	4.1	276.3
Maximum daily loss (kg/d)	9.8	4.51	2.4	32.0
Mean daily loss (kg/d)	4.0	1.89	1.3	14.8
Loss duration (d)	10.7	5.40	3	33

During the study, 501 HD were detected by farm staff. The most frequent HD were mastitis (n = 189), lameness (n = 144), metritis (n = 62), impaired general condition without identified cause (n = 58) and diarrhea (n = 25) (table 3-II-3). Other types of HD were very rare. Eleven rare and severe events were classified as accidents. For instance 2 cows were involuntarily blocked for 1 night under a stall, and 6 cows fell on concrete.

Table 3-II-3: number and severity of health disorders detected by farm staff concomitantly to a drop in milk production (between 5 d before and 5 d after) or independently (others)

Health disorder	Severity	HD independent	HD concomitant with a drop in milk production (% of the line)	Total
Mastitis	1	104	13 (11)	117
	2	44	13 (23)	57
	3	9	6 (40)	15
Lameness	2	100	6 (6)	106
	3	31	7 (18)	38
Metritis	1	27	0 (0)	27
	2	32	0 (0)	32
	3	3	0 (0)	3
Impaired general condition	1	17	20 (54)	37
	2	7	14 (67)	21
Diarrhea	-	16	9 (36)	25
Accident	-	9	2 (18)	11
Abortion	-	2	0 (0)	2
Digestive tract disorder	-	1	3 (75)	4
Retained placenta	2	3	0 (0)	3
Pericarditis	-	0	1 (100)	1
Pneumonia	-	2	0 (0)	2
Total		407	94 (19)	501

Results concerning the concomitancy between farm staff detections and DMP are displayed in table 3-II-3. Among the 501 HD detected by farm staff, 94 (19%) were concomitant with a major DMP. For the more frequent HD, the percentage of HD concomitant with DMP ranged from 0% for metritis to 67% for impaired general condition episodes with hyperthermia. The percentage of concomitancy increased with severity score of HD. On the other hand, 407 HD were not concomitant with DMP (table 3-II-3) and 263 DMP were not concomitant with HD.

Informative Value of Sensors' Detections

Due to excessive missing values from sensors, 59 DMP were excluded from this part of the analysis. Thus 288 DMP were included in the analysis for informative value calculation. With the chosen CUSUM setting, the specificity of the detections by the CUSUM on RT or on NA was 95.2%. Overall, 103 drops were detected by the CUSUM (Se = 35.8% ; 38.3% on farm 1 and 20% on farm 2). Mean total loss, mean daily loss, maximum daily loss and loss duration of detected vs. undetected drops were 48.0 vs. 35.1 kg ($P < 0.001$), 4.4 vs. 3.7 kg/d ($P < 0.01$), 11.1 vs. 8.9 kg/d ($P < 0.001$) and 11.2 vs. 10.4 d ($P > 0.05$) respectively. Sixty-four drops were detected by the CUSUM on RT or NA but not by farm staff, 30 were detected by farm staff but not by the CUSUM on RT or NA, 39 were detected by both methods and 155 were not detected by any method (table 3-II-4). Overall, 18.4% of independent HD, 29.2% of independent DMP and 56.5% of DMP concomitant with a HD detected by farm staff were detected by the CUSUM on RT or NA.

Table 3-II-4: results of the CUSUM on RT or NA (number of events) for the detection of the major drops in milk production (n = 288) and of the HD detected by farm staff (n = 416).

		Health disorder independent	Drop in milk independent	Health disorder concomitant with a drop in milk production	Total
Alarm by the CUSUM on RT or NA	Yes	64	64	39	167
	No	283	155	30	468
	Total	347	219	69	635
	Sensitivity (%) (95%CI)	18.4 (14.7-22.9)	29.2 (23.6-35.6)	56.5 (44.8-67.6)	26.3 (23.0-29.9)

Time-lag between Sensors' Detections and Drops in Milk

The time-lag between the first detection by the CUSUM on RT or NA and the beginning of the DMP for the detected DMP is displayed in figure 3-II-2. Overall, 60% of the detection by the CUSUM on RT or NA occurred on the same day as the DMP or on the 3 d before, 29% occurred 1 or 2 d after the drop and 11% occurred 3 or 4 d after the drop.

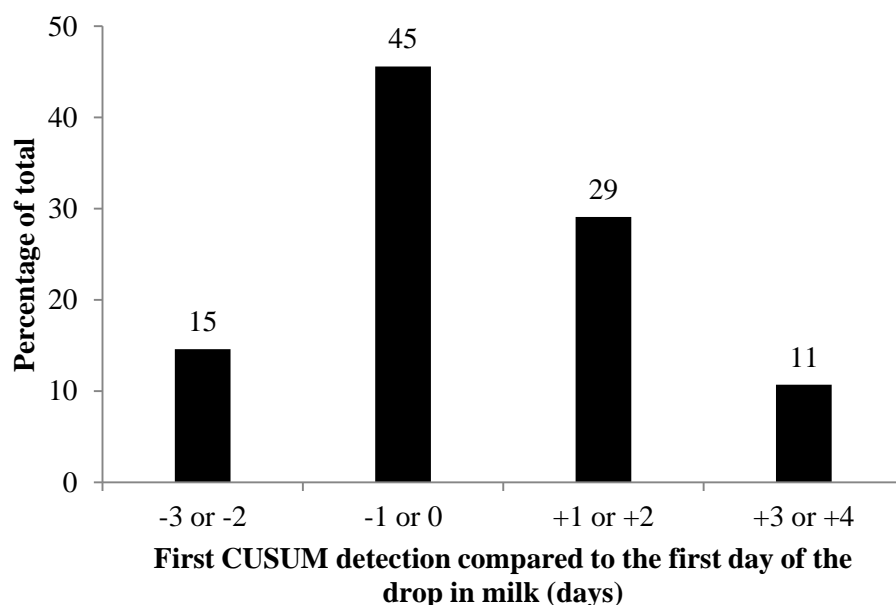


Figure 3-II-2: time-lag between the detection by the CUSUM on RT or NA (n = 103) and the first day of the drop in milk production.

6. Discussion

Our study demonstrates that major DMP were frequent and largely not concomitant with the detection of HD by farm staff. Indeed, among the 347 major DMP detected, 263 were not concomitant with a HD detected by farm staff. As the mean total loss (40.7 kg) and mean loss duration (10.7 d) of DMP included in the analysis were high, we can assume that the health status of the cows was truly affected around the DMP episodes, even if the exact nature of HD was unknown as no clinical signs were detected by farm staff. Moreover, individual DMI was measured daily on one third of the cows throughout the study. For these cows, drops in DMI were often detected a few hours before DMP (data not shown) using similar method, i.e. locally weighted polynomial and CUSUM on DMI residuals. This result reinforces the assumption that the health of the cows was truly affected during major DMP. The existence of HD impacting milk yield without any detection by farm staff is consistent with several studies that demonstrated that for HDs such as pneumonia (Timsit et al., 2011) or lameness (Leach et al., 2010) for instance, only a small number of sick animals were detected by visual appraisal by farm staff. In some cases, animals may also have coped with their disease, showing insufficient clinical signs for a detection and treatment by farm staff.

Among the 501 HD detected by farm staff, 407 were not concomitantly associated with a DMP. Several reasons may explain this result. First, some HD detected by farm staff may have no impact or a limited impact on milk yield, because they are mild in nature of their severity. For instance, in our study no metritis of severity 2 (endometritis) was concomitant with a DMP. Conversely, Bareille et al. (2003) found a significant effect of endometritis on milk yield but the model used in this study aimed at determining the mean effect of diseases on milk yield and not at detecting health events. Thus, a HD may provoke a decrease in milk yield which is not sudden or intense enough to be detected by our method. For instance, chronic lameness may provoke a very gradual DMP potentially not detected by the CUSUM on milk residuals. With our method, choosing the threshold above which a DMP had to be considered as major was arbitrary. We chose relatively specific thresholds, as half of the DMP initially detected by the CUSUM were excluded due to either insufficient amplitude or insufficient total relative loss. Another reason for the high number of HD without DMP may have been an early detection and treatment of HD by farm staff. In some cases, the treatment and the cure of the cow may have occurred before a major DMP was detectable. Despite this, table 3-II-3 shows that a large variety of HD were concomitant with DMP, which is consistent with other studies (Bareille et al., 2003, Fourichon et al., 1999).

This study is original as the CUSUM on RT or NA was evaluated using both HD detected by farm staff and major DMP as events to detect. The complementarity between farm staff detections and DMP detections for the evaluation of the CUSUM on RT or NA, or more generally for the evaluation of detection models based on sensor data, was highlighted. Table 3-II-4 shows that, only 416 and 288 events would have been considered detectable if visual appraisal or major DMP were used as the reference methods, respectively. If the detection by at least one of the 2 methods was the reference method, 635 events would have been considered detectable. Moreover, 64 alarms triggered by the CUSUM on RT or NA corresponded to major DMP, without concomitant HD detection by farm staff. In a study using only farm staff detection as a reference method to evaluate the algorithm on RT or NA, these alarms would have been considered as false positives and the specificity of the algorithm would have been underestimated. In commercial dairy farms, where individual daily milk yield is often not available, the apparent lack of specificity of an HD detection system could be due to the detection by the system of events such as DMP, not detected by farm staff (higher sensitivity of the HD detection system for the detection of DMP). On the other hand, some HD detected only by farm staff, without associated DMP, are expected to be detected by the monitoring device evaluated in this study, in order to enable an early diagnosis and treatment of diseased animals. Thus, it could be

relevant to combine DMP or detections by farm staff as a unique reference method to evaluate the sensitivity and the specificity of the CUSUM on RT or NA data. In this case, DMP or HD detected by farm staff would be considered as events relevant to detect. This method is different from the classical approach, where milk yield is used as an input variable to improve the performance of multivariate detection models aiming at detecting mastitis (Miekley et al., 2013) or lameness (Miekley et al., 2013, Van Hertem et al., 2013) for instance. Indeed, this classical approach requires having a good reference method to classify animals as sick or healthy, as the output of the detection model is compared with the health status of the cow determined by this reference method. In such models, major DMP may trigger detections considered as false positive if the reference method is not sensitive enough. It might be the case if visual appraisal by farm staff is used to detect mastitis or lameness. Conversely, our method considers a major DMP as a sign of disease by itself. Milk yield is a good indicator for HD detection: it is relatively inexpensive to collect, it may be impacted several days before the detection of an HD by farm staff (Bareille et al., 2003), and for some diseases such as Paratuberculosis, milk losses may be the only sign of the disease at an early stage (Beaudeau et al., 2007). Moreover, maximizing the detection of HD that impact milk yield is important, as these HD have a direct economic impact on farm profitability, due to milk losses. When available, other measures such as individual DMI or individual water intake could be used in addition to milk yield to determine whether animals are sick or healthy (Lukas et al., 2008, Miekley et al., 2013).

At the level of specificity calculated, the overall sensitivity of the algorithm on RT and NA data was low whatever the reference method used (table 3-II- 4), which is consistent with the results of a study by Clément et al. (chapter 3-I), even if a sensitivity over 55% was found in that study for HD impacting general condition. However choosing a more sensitive setting for the CUSUM on RT or NA would produce too much false alarms for practical and sustainable use of the device by the farmer (Clément et al., chapter 3-I).

Our results also demonstrate an early detection of DMP by the CUSUM on RT or NA, with 60% of the detections occurring between the first day of the DMP and 3 d before, and 30% of the detections occurring one or two days after the beginning of the DMP. Conversely, we can assume that a detection model using milk yield as an input data would have detected DMP after a period of at least one or two days. Moreover, for the 37 DMP detected by both farm staff and the CUSUM on RT or NA, the precocity of the detections by the CUSUM tends to be higher than the detections by farm staff (on average detection 12 h before the detection by farm staff, data not shown).

7. Conclusions

Many major DMP not concomitant with a detection of HD by farm staff were detected during lactation (0.7 per lactation). These DMP can be considered as a clinical sign of HD and may trigger alarms by the monitoring device (0.2 per lactation). Thus, there is a risk to underestimate the specificity of the monitoring device evaluated if these DMP are not considered in the analysis as events to detect. In our study, these drops corresponding to important milk losses were sometimes associated with alarms triggered by a CUSUM on RT or NA. Thus we can assume that these events correspond to true alterations of the health status of the cows, deemed to be detected. A detailed clinical examination during these DMP should enable to determine the nature of the HD affecting the cow. Early detection of some HD at subclinical stages would be possible but there is also a risk to signal to the farmer some self-healing cases that would otherwise never have come forward. Further work is needed to determine optimal threshold to declare a DMP as major or not.

References

- Bareille, N., F. Beaudeau, S. Billon, A. Robert, and P. Faverdin. 2003. Effects of health disorders on feed intake and milk production in dairy cows. *Livest Prod Sci* 83:53-62.
- Beaudeau, F., M. Belliard, A. Joly, and H. Seegers. 2007. Reduction in milk yield associated with *Mycobacterium avium* subspecies paratuberculosis (Map) infection in dairy cows. *Vet Res* 38:625-634.
- Bewley, J. M. 2010. Precision dairy farming: Advanced analysis solutions for future profitability. Accessed July, 20, 2014. <http://www.precisiondairy2010.com/proceedings/s1bewley.pdf>.
- Cleveland, W. and E. Grosse. 1991. Computational methods for local regression. *Stat Comput* 1:47-62.
- de Koning, C. J. A. M. 2010. Automatic milking-Common practice on dairy farms in Proc. First North. Am. Conf. Precision Dairy Management, Toronto, Canada, Pages 52-67.
- Fabian, J., R. A. Laven, and H. R. Whay. 2014. The prevalence of lameness on New Zealand dairy farms: A comparison of farmer estimate and locomotion scoring. *Vet J* 201:31-38.
- Fourichon, C., H. Seegers, N. Bareille, and F. Beaudeau. 1999. Effects of disease on milk production in the dairy cow: a review. *Prev Vet Med* 41:1-35.
- Hogeveen, H., C. Kamphuis, W. Steeneveld, and H. Mollenhorst. 2010. Sensors and clinical mastitis--the quest for the perfect alert. *Sensors (Basel)* 10:7991-8009.
- Huybrechts, T., K. Mertens, J. De Baerdemaeker, B. De Ketelaere, and W. Saeys. 2014. Early warnings from automatic milk yield monitoring with online synergistic control. *J Dairy Sci* 97:3371-3381.
- Kamphuis, C., B. DelaRue, C. R. Burke, and J. Jago. 2012. Field evaluation of 2 collar-mounted activity meters for detecting cows in estrus on a large pasture-grazed dairy farm. *J Dairy Sci* 95:3045-3056.
- Leach, K. A., H. R. Whay, C. M. Maggs, Z. E. Barker, E. S. Paul, A. K. Bell, and D. C. Main. 2010. Working towards a reduction in cattle lameness: 1. Understanding barriers to lameness control on dairy farms. *Res Vet Sci* 89:311-317.

- Lukas, J. M., J. K. Reneau, and J. G. Linn. 2008. Water intake and dry matter intake changes as a feeding management tool and indicator of health and estrus status in dairy cows. *J Dairy Sci* 91:3385-3394.
- Lukas, J. M., J. K. Reneau, R. Wallace, D. Hawkins, and C. Munoz-Zanzi. 2009. A novel method of analyzing daily milk production and electrical conductivity to predict disease onset. *J Dairy Sci* 92:5964-5976.
- Luo, Y., Z. Li, and Z. Wang. 2009. Adaptive CUSUM control chart with variable sampling intervals. *Comput. Stat. Data Anal.* 53:2693-2701.
- Miekley, B., I. Traulsen, and J. Krieter. 2012. Detection of mastitis and lameness in dairy cows using wavelet analysis. *Livest Sci* 148:227-236.
- Miekley, B., E. Stamer, I. Traulsen, and J. Krieter. 2013. Implementation of multivariate cumulative sum control charts in mastitis and lameness monitoring. *J Dairy Sci* 96:5723-5733.
- Radostits, O.M., C.C. Gay, K.W. Hinchcliff, and P. D. Constable. 2007. *Veterinary Medicine: A textbook of the diseases of cattle, horses, sheep, pigs, and goats*. 10th ed. Saunders Elsevier, New York
- Reith, S. and S. Hoy. 2012. Relationship between daily rumination time and estrus of dairy cows. *J Dairy Sci* 95:6416-6420.
- Rutten, C. J., A. G. Velthuis, W. Steeneveld, and H. Hogeveen. 2013. Invited review: sensors to support health management on dairy farms. *J Dairy Sci* 96:1928-1952.
- Schirmann, K., M. A. von Keyserlingk, D. M. Weary, D. M. Veira, and W. Heuwieser. 2009. Technical note: Validation of a system for monitoring rumination in dairy cows. *J Dairy Sci* 92:6052-6055.
- Schulze, C., J. Spilke, and W. Lehner. 2007. Data modeling for Precision Dairy Farming within the competitive field of operational and analytical tasks. *Comput Electron Agr* 59:39-55.
- Timsit, E., N. Bareille, H. Seegers, A. Lehebel, and S. Assie. 2011. Visually undetected fever episodes in newly received beef bulls at a fattening operation: occurrence, duration, and impact on performance. *J Anim Sci* 89:4272-4280.

Van Hertem, T., E. Maltz, A. Antler, C. E. Romanini, S. Viazzi, C. Bahr, A. Schlageter-Tello, C. Lokhorst, D. Berckmans, and I. Halachmi. 2013. Lameness detection based on multivariate continuous sensing of milk yield, rumination, and neck activity. *J Dairy Sci* 96:4286-4298.

Partie III. Evaluation d'alertes de santé générées chez des vaches laitières à partir de données d'activité et de rumination issues de capteurs, par rapport à un examen clinique vétérinaire systématique.

1. Résumé

Les objectifs de cette étude étaient (i) d'évaluer la sensibilité et la spécificité d'alertes de santé générées à partir des durées de rumination et de l'activité chez des vaches laitières, par rapport à leur Examen Clinique Vétérinaire (ECV) systématique et détaillé (ii) de juger de l'intérêt de la réalisation d'ECV pour évaluer la valeur informative d'alertes de santé pour des systèmes de monitoring de la santé.

L'étude s'est déroulée durant 6 mois dans une ferme laitière de l'ouest de la France comptant 90 vaches Holstein en moyenne. Chaque matin, des alertes CUSUM sensibles (sensibilité = 66% ; spécificité = 41%) étaient appliquées sur les données brutes d'activité et de rumination issues d'un capteur et recueillies au cours de la traite. Une liste de vaches en alerte, ayant présenté une chute d'activité ou de rumination dans les 4 derniers jours, était générée quotidiennement, ainsi qu'une liste de vaches n'ayant pas présenté de chutes d'activité ou de rumination dans les 7 derniers jours (vaches non en alerte). Toutes les vaches en alerte et certaines vaches tirées au sort parmi celles non en alerte subissaient ensuite un ECV détaillé afin de déterminer leur état de santé. Les vaches signalées par l'éleveur comme présentant une anomalie (baisse d'appétit, de production laitière, signe clinique...) subissaient également un ECV. A posteriori, des CUSUM optimisés ont été ré-appliqués sur les données avec de nouveaux paramètres permettant d'obtenir une spécificité suffisante (79%), et la sensibilité des alertes de santé pour la détection des principaux troubles de santé rencontrés a été calculée.

Au total, 276 examens cliniques vétérinaires ont été réalisés sur 103 vaches. Parmi les 240 ECV inclus dans l'analyse, 184 (76,3% des ECV) étaient négatifs (absence de maladie diagnostiquée) et 56 ECV étaient positifs. Les principales maladies diagnostiquées étaient les boiteries (n = 19), les mammites (n = 17) et les métrites (n = 13). L'éleveur avait signalé spontanément une anomalie chez 4% des animaux ECV négatifs et chez 37% des animaux ECV positifs. Pour une spécificité de 79%, 32% de l'ensemble des maladies diagnostiquées donnaient lieu à une alerte de santé dans les 4 jours précédant le diagnostic de la maladie, soit respectivement 11% des boiteries, 35% des mammites et 46% des métrites.

En raison de la faible incidence de troubles de santé dans l'élevage, trop peu de vaches ont été diagnostiquées malades pour pouvoir estimer avec une bonne précision la sensibilité des alertes de santé utilisées. Cependant, le design de l'étude a permis de mettre en évidence que certaines maladies comme les boiteries ou les métrites étaient sous-détectées par les éleveurs. Ce design original pourrait également être utilisé avec pertinence pour calculer la sensibilité et la spécificité d'alertes de santé dans des contextes d'élevage où l'incidence des troubles de santé est plus élevée, comme par exemple en élevages de jeunes bovins en engraissement.

La comparaison des sensibilités et spécificités obtenues avec celles obtenues dans une autre étude utilisant la détection par les éleveurs comme méthode de référence, montre que ces valeurs ne sont pas estimées beaucoup plus précisément lorsqu'un ECV systématique des vaches est effectué.

2. Introduction

Depuis les années 2000, de nombreux systèmes de monitoring ont été développés pour assister les éleveurs pour la détection des vèlages, des chaleurs ou de certaines maladies (boiteries, mammites, maladies métaboliques par exemple (Rutten et al., 2013)). Ces systèmes génèrent des alertes à destination des éleveurs, à partir de mesures physiques (température, activité) ou chimiques (pH ruminal) faites en continu sur les animaux à l'aide de capteurs. Un paramétrage adéquat des alertes est nécessaire au moment de leur développement, pour permettre une utilisation réelle et optimale de l'appareil (cf chapitre 3.I et 3.II). Le calcul de statistiques comme la sensibilité et la spécificité permettent de paramétrer les alertes de manière optimale. Pour le calcul de ces statistiques, il faut pouvoir déterminer avec le plus d'exactitude possible le statut de santé, malade versus non malade, des animaux. Pour cela on utilise une méthode de référence. Dans les chapitres précédents, nous avons utilisé comme méthodes de référence la détection par observation visuelle de l'éleveur (chapitre 3.I) et le repérage de chutes majeures de production laitière (chapitre 3.II) pour déterminer l'état de santé des animaux. Cependant, plusieurs études ont montré que la détection des troubles de santé par observation visuelle des éleveurs était peu sensible pour la détection de certaines maladies comme les troubles respiratoires en atelier d'engraissement de jeunes bovins (Timsit et al., 2011b) ou les boiteries chez les vaches laitières (Fabian et al., 2014). Certains troubles de santé non détectés par les éleveurs peuvent donc déclencher des alertes de santé, ce qui peut conduire à sous-estimer la spécificité des alertes et à une mauvaise estimation de la sensibilité lorsque la détection par l'éleveur est considérée comme

la méthode de référence. Dans le chapitre 3.II, nous avons eu des difficultés à déterminer un seuil au-delà duquel une chute de production laitière devait être considérée comme suffisamment importante pour être détectée, rendant cette méthode de référence discutable. De plus, cette méthode de référence ne permet pas de déterminer de quelle maladie sont atteints les animaux.

Lorsqu'ils font face à des troubles de santé qu'ils ne peuvent diagnostiquer et/ou traiter eux-mêmes, les éleveurs font généralement appel à leur vétérinaire. L'examen clinique vétérinaire (ECV) permet de diagnostiquer des maladies à partir des signes cliniques observés chez les vaches. Il permet donc de déterminer l'état de santé des animaux, probablement avec une bonne fiabilité. En effet, alors que l'éleveur se base principalement sur l'observation, l'inspection et les données de production pour détecter les maladies, les vétérinaires peuvent en plus avoir recours à la palpation, la percussion, l'auscultation des différents appareils ainsi qu'à des examens complémentaires. Cette méthode de référence a été utilisée dans quelques études chez la vache laitière pour diagnostiquer un seul trouble de santé comme l'acétonémie (Gillund et al., 2001) ou les métrites (Sepulveda-Varas et al.) ou déterminer plus généralement l'état de santé des animaux (Steen, 2001). Chez les bovins allaitants, cette méthode a été utilisée pour évaluer la précocité des alertes fournies par un thermobolus ruminal (Timsit et al., 2011a) pour la détection des troubles respiratoires par exemple. A notre connaissance, très peu d'auteurs ont utilisé l'ECV pour l'évaluation d'un outil de monitoring pouvant détecter différents troubles de santé chez les vaches laitières.

Les objectifs de cette étude étaient (i) d'évaluer la sensibilité et la spécificité d'alertes de santé générées à partir de l'activité et des durées de rumination chez des vaches laitières, par rapport à leur ECV systématique et détaillé (ii) de juger de l'intérêt de la réalisation d'ECV pour évaluer la valeur informative d'alertes de santé générées par des systèmes de monitoring de la santé.

3. Matériel et méthodes

3.1 Principe et design de l'étude

Pour pouvoir déterminer la valeur informative des alertes générées à partir des données brutes d'activité et de rumination fournies par Heatime-HR® (chapitre 3.I, Schirmann et al., 2009) , et calculer la Sensibilité (Se) et la Spécificité (Sp) des alertes en comparaison avec un examen clinique vétérinaire (ECV), il fallait pouvoir remplir le tableau suivant (tableau 3-III-1).

Tableau 3-III-1 : principe du calcul de la sensibilité (Se) et de la spécificité (Sp) des alertes de santé (a, b, c, d = effectifs).

		Examen clinique vétérinaire		
		Malade	Non malade	
Algorithmes de détection des baisses d'activité/ rumination	Alerte de santé	a	b	$Se = \frac{a}{a + c} * 100$ $Sp = \frac{d}{d + b} * 100$
	Non en alerte de santé	c	d	

La sensibilité est le pourcentage de vaches malades en alerte de santé (détectées par les algorithmes de détection des baisses d'activité ou de rumination). La spécificité est le pourcentage de vaches non malades non en alerte de santé (non détectées par les algorithmes) (tableau 3-III-1).

3.2 Echantillon d'étude

3.2.1 Critères de choix du troupeau

Les données de la bibliographie ne permettaient pas de faire des hypothèses *a priori* concernant le remplissage du tableau de contingence pour le calcul de la sensibilité et de la spécificité des alertes de santé. En particulier, aucun calcul de la taille d'échantillon nécessaire pour répondre aux objectifs de l'étude n'étant possible, nous avons travaillé sur un échantillon qui s'adaptait le mieux aux contraintes de l'étude :

- Le(s) élevage(s) devai(en)t être équipé(s) du dispositif Heatime-HR®, qui fournissait les données brutes d'activité et de rumination.
- Les examens cliniques vétérinaires et le recueil des données Heatime-HR® devaient être effectués quotidiennement sur la période d'étude.
- Le budget disponible pour cette étude permettait le recrutement d'un vétérinaire pendant 6 mois pour effectuer l'ensemble des ECV.

Compte tenu de la faible densité d'élevages équipés, l'étude a été effectuée dans un troupeau d'une ferme commerciale (Sucé sur Erdre, Loire-Atlantique) répondant à l'ensemble de ces critères, durant 2 périodes successives. Le protocole de l'étude a été appliqué entre le 01/11/2012 et le 24/01/2013 (période 1) et entre le 27/11/2013 et le 14/02/2014 (période 2). Ces périodes ont été choisies pour maximiser les chances d'examiner des vaches malades, car elles correspondaient aux périodes où la plupart des vaches étaient en début de lactation (vêlages entre fin septembre et fin décembre), c'est-à-dire lorsque l'incidence des troubles de santé est maximale (Gernand et al., 2012). L'étude a été effectuée par un premier vétérinaire durant la période 1 et par un second vétérinaire durant la période 2.

3.2.2 Animaux et conditions d'élevage

Dans cet élevage, en moyenne 90 vaches Prim'Holstein étaient présentes. Les vaches étaient traitées deux fois par jour à 7h et 17h. La ferme était équipée en logettes ainsi qu'en cornadis. Une ration semi-complète était distribuée après la traite du matin. La ration de base durant les deux périodes était composée de 16kg de matière sèche ($\frac{3}{4}$ de maïs et $\frac{1}{4}$ d'ensilage d'herbe) plus 4kg de correcteur azoté. Le matin, les vaches étaient bloquées aux cornadis entre la fin de leur passage en salle de traite et environ une heure après la fin de la traite, ce qui permettait de réaliser les ECV durant 1h30 environ. Les données d'activité ou de rumination étaient transférées à l'unité centrale Heatime-HR® au moins une fois par jour lors du passage des vaches sous une des 2 antennes infrarouge du bâtiment (une à la sortie de la salle de traite et une au-dessus d'un abreuvoir du box des vaches tarées). La récupération des données n'était pas systématique pendant la période de tarissement, certaines vaches tarées restant au pâturage. Les vaches étaient très généralement rentrées en stabulation durant les 2 périodes de l'étude. Cependant elles pouvaient être mises à l'herbe exceptionnellement la journée quand les conditions climatiques le permettaient.

Sur la période 1, le troupeau laitier était composé en moyenne de 90 vaches laitières dont 27 primipares et 62 multipares. Le rang de lactation était en moyenne de 2,7 et le stade de lactation de 6 mois. Sur la période 2, il y avait dans le troupeau 104 vaches laitières dont 45 primipares et 59 multipares. Le rang de lactation moyen était de 2,6 et le stade de lactation moyen était de 7 mois.

3.3 Algorithmes initiaux et production des alertes de santé

Dans cette étude, le statut des animaux du troupeau vis-à-vis d'alertes de santé a été déterminé dans un premier temps, puis un ECV d'animaux en alerte et non en alerte a été effectué (cf tableau 3-III-1). En effet, un point sensible identifié préalablement à la réalisation du protocole était la difficulté de recruter suffisamment d'animaux présentant des troubles de santé pour le calcul de la sensibilité des détections. Des algorithmes sensibles ont donc été privilégiés, en faisant l'hypothèse que les vaches présentant des chutes de rumination ou d'activité avaient plus de risques de présenter des troubles de santé. Les paramètres des algorithmes initiaux ont été fixés suite à une précédente étude effectuée à la ferme expérimentale des Trinottières (49) car ils permettaient d'obtenir une sensibilité élevée des alertes de santé par rapport à la détection par les éleveurs.

Les algorithmes développés pour cette étude utilisaient un test *cumulative sum* (CUSUM) pour détecter des chutes de temps de rumination ou d'activité (Luo et al., 2009). Le CUSUM additionne les différences, appelées erreurs, entre un temps de rumination ou une activité de référence et le temps de rumination ou l'activité observés, lorsque cette différence est positive.

Un test CUSUM proche de celui utilisé dans le chapitre 3.I a été utilisé pour produire les alertes de santé (Algorithmes initiaux), avec des modifications majeures concernant notamment le calcul de l'activité et de la durée de rumination de référence (tableau 3-III-2). L'algorithme utilisé était une première version assez simple de CUSUM, moins raffinée que l'algorithme qui a été développé dans la suite des travaux menés (chapitre 3.I).

Tableau 3-III-2 : paramètres utilisés pour les algorithmes CUSUM initiaux et optimisés (k=constante du CUSUM, h=limite de décision du CUSUM, cf chapitre 3.I).

	Chapitre 3.III-Algorithmes initiaux	Chapitre 3.I-Algorithmes optimisés
Fenêtre de calcul de la référence	5	7
Calcul de la référence	Moyenne de toutes les valeurs des 5j précédents	Moyenne transversale (valeurs à la même heure de la journée)
k du CUSUM	0,5	0,7
h rumination	5	100
h activité	4,5	75

Le logiciel SAS 9.2 (SAS Institute Inc., Cary, NC) a été utilisé pour le développement et l'application des algorithmes CUSUM.

Les données d'activité et de rumination étaient récupérées sur une clé USB sur l'unité centrale Heatime-HR® de l'élevage pendant la traite du matin, et les algorithmes CUSUM étaient appliqués afin de générer 2 types d'alertes de santé :

- Une alerte de rumination correspondant à une diminution du temps de rumination dans les 4 derniers jours.
- Une alerte d'activité correspondant à une diminution de l'activité dans les 4 derniers jours.

Ceci permettait d'obtenir une liste de vaches qui étaient soit en alerte de rumination, soit en alerte d'activité soit les deux à la fois. Une liste de vaches n'ayant pas présenté d'alerte de rumination ou d'activité dans les 7 jours précédents, considérées comme non en alerte, était également générée chaque matin. Toutes les vaches en alerte et certaines vaches non en alerte subissaient un ECV afin de déterminer leur état de santé.

3.4 Examen clinique vétérinaire

3.4.1 Critères d'inclusion et d'exclusion et choix des vaches à examiner

Les critères d'inclusion ou d'exclusion des animaux à examiner étaient les suivants :

- Les vaches devaient être équipées d'un collier Heatime-HR®.
- Elles devaient figurer soit dans la liste des vaches en alerte soit dans la liste des vaches non en alerte du jour.
- Les vaches repérées malades par l'éleveur par observation visuelle (baisse de production laitière, d'ingestion, suspicion de maladie spécifique) subissaient un ECV au même titre que les autres vaches.
- Elles ne devaient pas être en chaleur le jour de l'ECV car il a été montré que les chaleurs s'accompagnent fréquemment d'une diminution significative des durées de rumination (Reith et Hoy, 2012).
- Elles ne devaient pas avoir vêlé dans les 4 jours précédant une alerte de santé car une baisse de rumination et d'activité pouvant aller jusqu'à 72h post-vêlage est fréquemment observée de manière physiologique (Schirmann et al., 2013).
- Les vaches non en alerte ne devaient pas avoir eu d'alertes de santé dans les 7 derniers jours.

Compte tenu de la faible fréquence des troubles de santé en élevage laitier, il était attendu *a priori* que suffisamment de vaches non malades (effectifs b + d, tableau 3-III-1) soient disponibles pour le calcul de la spécificité mais que le recrutement de suffisamment de vaches malades (effectifs a + c, tableau 3-III-1) soit plus difficile. Toutes les vaches en alerte (effectifs a + b, tableau 3-III-1) ainsi que les vaches malades repérées par l'éleveur étaient donc examinées en priorité en faisant l'hypothèse qu'elles présentaient plus de chances d'être classées malades à l'issue de ECV que les vaches non en alerte. Des vaches non en alerte étaient examinées les jours où il n'y avait pas d'alertes ou lorsque qu'il restait du temps après l'examen clinique de toutes les autres vaches. Les vaches examinées étaient tirées au sort aléatoirement parmi la liste de vaches non en alerte vérifiant les critères d'inclusion.

3.4.2 Réalisation des examens cliniques vétérinaires

Les ECV ont été réalisés par un vétérinaire durant la période 1 et par un autre vétérinaire durant la période 2 (cf 3.2.1). Une période de formation de 10 jours a eu lieu en début de la deuxième période de l'étude pour assurer une notation homogène des formulaires utilisés pour la réalisation de l'ECV (Annexe 3-III-1) entre les 2 périodes d'étude. L'ECV était un examen classique général suivi d'un examen clinique spécial par appareil. Différents appareils étaient examinés systématiquement : l'appareil cardio-vasculaire, l'appareil respiratoire, l'appareil digestif, l'appareil locomoteur et l'appareil génital.

Trois grilles cliniques (Annexe 3-III-2) ont été utilisées lors de ces examens cliniques, concernant :

- Les boiteries (modifiée d'après Sprecher et al., (1997))
- Les métrites (modifiée d'après Sheldon et al. (2006)) étaient diagnostiquées à l'aide d'un vaginoscope. L'examen au vaginoscope n'était pas effectué sur les vaches à moins de 15 jours post-partum en l'absence de signes cliniques généraux. En effet, des pertes mucopurulentes physiologiques peuvent être observées jusqu'à 10 jours après le vêlage. Un examen à partir de 15 jours permet d'éviter alors un diagnostic de métrite à tort.
- Les mammites. Contrairement au reste de l'ECV, la détection des mammites était réalisée exclusivement par l'éleveur par palpation de la mamelle et examen des premiers jets en cas de suspicion de mammite. En effet, un produit de post-trempage était appliqué sur chaque animal à la traite ce qui ne permettait pas aux vétérinaires d'examiner systématiquement l'aspect du lait lors de l'ECV.

Le formulaire récapitulatif de l'ECV permettait de renseigner librement les observations du vétérinaire afin d'être le plus précis possible dans la description des signes cliniques observés. Le tableau 3-III-3 récapitule les principaux paramètres observés ou mesurés lors de l'ECV.

Nous avons fait l'hypothèse que l'acétonémie subclinique pouvait avoir un impact sur durées de rumination et donc sur les alertes de santé. Un dosage de la concentration sanguine en bêtahydroxybutyrate était donc systématiquement réalisé lors de l'ECV afin de caractériser le statut des vaches vis-à-vis de l'acétonémie. Pour cela, un prélèvement de sang était effectué sur tube à EDTA ponctionné à la veine coccygienne grâce à un système Vacutainer® stérile. Les dosages ont été faits à partir d'un appareil portable, Precision® Xceed™. Le principe de

fonctionnement de l'appareil est de déposer une goutte de sang sur une bandelette introduite dans l'appareil. Après quelques secondes, la concentration en bêtahydroxybutyrate de l'échantillon s'affiche sur l'appareil. Les valeurs obtenues sont bien corrélées aux concentrations sanguines en bêtahydroxybutyrate ($r = 0,95$) (Iwersen et al., 2009).

Tableau 3-III-3 : principaux paramètres observés lors de l'ECV.

Appareil	Paramètres observés ou mesurés
Etat général	<ul style="list-style-type: none"> • Habitus • Température rectale (°C) • Fréquence cardiaque (bpm) • Fréquence respiratoire (mpm) • Fréquence des contractions ruminales (mpm) • Aspect des muqueuses • Note d'état corporel (grille de 0 à 5)
Cardio-vasculaire	<ul style="list-style-type: none"> • Auscultation cardiaque (fréquence, rythme, intensité) • Présence ou absence d'œdème • Etat de réplétion des veines jugulaires
Respiratoire	<ul style="list-style-type: none"> • Observation de la courbe respiratoire • Signalement de bruits à distance • Auscultation pulmonaire
Digestif	<ul style="list-style-type: none"> • Auscultation de l'appareil digestif (rumen) • Consistance/couleur/aspect des bouses • Consistance du contenu ruminal
Locomoteur	<ul style="list-style-type: none"> • Score boiterie (grille annexe 3-III-2)
Génital	<ul style="list-style-type: none"> • Palpation transrectale de l'appareil génital • Vaginoscope, score métrite (grille annexe 3-III-2) • Score mammite (grille annexe 3-III-2)
Métabolique	<ul style="list-style-type: none"> • Dosage des corps cétoniques sanguins

Suite à cet examen, si un dysfonctionnement était diagnostiqué, l'animal était classé comme malade. Dans le cas contraire, il était considéré comme non malade. Les vaches en alerte de

santé classées non malades lors d'un premier ECV subissaient un examen clinique le lendemain et le surlendemain si elles étaient toujours en alerte de santé les 2 jours suivants, dans l'hypothèse où les alertes de santé seraient précoces par rapport à la survenue des signes cliniques. Après ces 2 jours, elles n'étaient plus examinées.

3.5 Stratégie d'analyse

Les objectifs de cette étude étaient (i) d'évaluer la sensibilité et la spécificité d'alertes de santé générées à partir de l'activité et des durées de rumination chez des vaches laitières, par rapport à leur examen clinique vétérinaire systématique et détaillé (ii) de juger de l'intérêt de la réalisation d'ECV pour évaluer la valeur informative d'alertes de santé. Compte tenu de ces objectifs, nous avons choisi de calculer la sensibilité et la spécificité non pas des algorithmes appliqués au cours de l'étude mais des algorithmes optimisés (tableau 3-III-2), développés après le début de cette étude, mais qui ont été évalués en utilisant la détection par l'éleveur comme méthode de référence dans le chapitre 3.I. Cela permettait, les algorithmes évalués étant identiques dans les chapitres 3.I et 3.III, de juger de la pertinence d'utiliser l'ECV comme méthode de référence. En particulier, il était attendu que la spécificité des alertes de santé soit meilleure dans le chapitre 3.III, les vétérinaires classant potentiellement davantage d'animaux en alerte de santé malades (vrais positifs) que les éleveurs (mêmes animaux considérés comme des faux positifs en raison de la plus faible sensibilité attendue de l'éleveur).

3.5.1 Exposition aux alertes de santé optimisées

A posteriori, des algorithmes CUSUM identiques à ceux du chapitre 3.I (tableau 3-III-2, Algorithmes optimisés) ont donc été appliqués sur les données d'activité et de rumination, afin de déterminer la valeur informative des alertes générées par ces algorithmes optimisés. Les vaches étaient considérées en alerte le jour j si elles déclenchaient une alerte de santé (CUSUM sur l'activité ou la rumination) entre $j-4$ jours et $j+1$ jour. Sinon, elles étaient considérées comme non en alerte. Si plus de 33% des données d'activité ou de rumination étaient manquantes sur la fenêtre de 5 jours, le jour j considéré était exclu de l'analyse.

3.5.2 Analyse du statut de santé

Pour chaque examen clinique, l'ensemble des signes cliniques observés sur l'animal a été pris en compte et, lorsque c'était possible, un diagnostic de maladie a été posé. Par exemple la maladie « suspicion de réticulo-péricardite » était diagnostiquée lorsque les signes cliniques « hyperthermie », « baisse d'appétit », « tachycardie » et « bruits liquidiens à l'auscultation cardiaque » étaient observés. Des signes cliniques majeurs (par exemple hyperthermie sans cause identifiée) pouvaient être conservés seuls si aucune maladie n'était diagnostiquée.

Lorsque des ECV étaient réalisés à 1 jour d'intervalle sur la même vache (en cas d'alerte de santé initiale de longue durée), seul le premier ECV était conservé pour l'analyse lorsque les ECV successifs étaient négatifs (pas de maladie), ou seul le premier ECV positif était conservé dans le cas contraire. Lorsqu'un ECV était réalisé moins de 8 jours après le diagnostic d'une maladie, cet ECV était également ignoré lors de l'analyse, l'animal étant considéré comme non guéri.

Les vaches étaient considérées en acétonémie subclinique si elles présentaient une concentration sanguine en bêtahydroxybutyrate supérieure à 1,2 mmol/L (Iwersen et al., 2009).

3.5.3 Calcul des valeurs informatives des alertes de santé optimisées

Pour le calcul des valeurs informatives (sensibilité et spécificité), l'unité statistique était l'ECV. Les résultats des ECV conservés pour l'analyse ont été utilisés pour calculer la sensibilité et la spécificité des alertes en remplissant un tableau de contingence similaire au tableau 3-III-1 (cf 3.III.3.1). Pour l'analyse de la sensibilité globale des détections (tous troubles confondus), ont été considérées comme malades toutes les vaches pour lesquelles une maladie avait été diagnostiquée lors de l'ECV. Néanmoins, afin de pouvoir comparer les résultats avec le chapitre 3.I, les boiteries légères ont été comptabilisées dans la catégorie non malades. Les ECV pour lesquels la seule anomalie était une acétonémie subclinique ont également été considérés négatifs (vache classée non malade). La sensibilité tous troubles de santé confondus et la spécificité ont été calculées en utilisant les formules du tableau 3-III-1 (cf 3.III.3.1). Les alertes de santé optimisées générées chez des vaches classées malades à l'issue de l'ECV mais non détectées malades par l'éleveur ont été comptabilisées.

La sensibilité des alertes de santé optimisées a ensuite été calculée pour les 3 maladies les plus fréquemment rencontrées (les boiteries, les mammites et les métrites). Pour un trouble donné, la sensibilité était également calculée en fonction de la sévérité du trouble, en utilisant les catégories des grilles utilisées pour la notation des troubles (annexe 3-III-2). Pour chaque sensibilité, un intervalle de confiance à 95% a été calculé.

4. Résultats

4.1 Statistiques descriptives

Sur les 2 périodes d'étude, 276 examens cliniques ont été réalisés sur 103 vaches. La distribution des alertes de santé initiales était la suivante : 22% des ECV étaient réalisés suite à une alerte sur la rumination seule, 18% suite à une alerte sur l'activité seule, 5% suite à une alerte sur l'activité et la rumination, et 55% suite à une absence d'alerte de santé. Après exclusion des ECV réalisés à un jour d'intervalle ($n = 28$), des ECV réalisés moins de 8 jours après un ECV positif ($n = 3$), et des ECV pour lesquels plus de 33% des données d'activité ou de rumination étaient manquantes dans la fenêtre de 5 jours ($n = 5$), 240 ECV ont été conservés pour l'analyse. Soixante-cinq vaches de l'échantillon ont été examinées plus d'une fois, 38 ont été examinées une seule fois. Sur l'ensemble des ECV inclus dans l'analyse, 80 vaches étaient primipares et 160 étaient multipares au moment de l'examen clinique. Le rang moyen de lactation était de 2,4. Le stade de lactation moyen était de 186 jours. Parmi les 240 ECV inclus dans l'analyse, 56 étaient positifs (présence d'une maladie) et 184 étaient négatifs (absence de maladie). Les différentes maladies diagnostiquées ainsi que le signalement ou non d'une anomalie par l'éleveur concernant l'animal examiné (baisse de production laitière, d'ingestion, suspicion de maladie...) sont présentés dans le tableau 3-III-4.

Tableau 3-III-4 : description des maladies et anomalies diagnostiquées lors des examens cliniques vétérinaires (ECV).

Maladie (résultat de l'ECV)	Anomalie signalée par l'éleveur (% de la ligne)	Pas d'anomalie signalée (% de la ligne)	Nombre total	% des ECV
Absence de maladie	7 (4)	177 (96)	184	76,3
Aucune anomalie à l'ECV	4	130	134	55,6
Boiterie légère	1	13	14	5,8
Acétonémie subclinique	2	34	36	14,9
Présence de maladie	21 (37)	36 (63)	57	23,7
Boiterie modérée	1	14	15	6,2
Boiterie sévère	0	4	4	1,7
Total boiteries	1 (5)	18 (95)	19	7,9
Mammite légère	6	0	6	2,5
Mammite modérée	8	0	8	3,3
Mammite sévère	3	0	3	1,2
Total mammites	17 (100)	0 (0)	17	7,1
Métrite	2 (15)	11 (85)	13	5,4
Suspicion de péricardite par corps étranger	0	1	1	0,4
Déplacement de caillette à gauche	0	1	1	0,4
Autre trouble digestif	0	2	2	0,8
Hyperthermie sans cause identifiée	0	2	2	0,8
Trouble respiratoire	0	1	1	0,4
Abcès hanche postérieur gauche	1	0	1	0,4
Total	28	213	241	100,0

Une majorité des ECV (76,3%) ne donnait pas lieu à un diagnostic de maladie. Les principaux troubles de santé diagnostiqués étaient les boiteries (n=19), les mammites (n=17) et les métrites (n=13). Lors d'un ECV seulement, des signes cliniques pouvant être attribués à des maladies différentes ont été diagnostiqués sur la même vache : une vache présentait à la fois une mammite modérée et une boiterie modérée.

Dans 4% des cas où l'ECV était négatif (absence de maladie), l'éleveur avait signalé une anomalie (tableau 3-III-4). Dans 37% des cas où l'ECV était positif (présence de maladie), l'éleveur avait signalé une anomalie chez l'animal. De même, l'éleveur avait signalé une anomalie dans 5% des cas de boiterie, et dans 15% des cas de métrite. Cent pourcent des mammites donnaient lieu à un signalement par l'éleveur puisqu'il assurait leur détection (présence d'un produit de post-trempage).

4.2 Valeur informative des alertes de santé optimisées

La spécificité calculée des alertes de santé élaborées était de 79% ($IC_{95\%} = 72-84$). Parallèlement, la sensibilité des alertes optimisées tous troubles de santé confondus était de 32% ($IC_{95\%} = 20-46$). La sensibilité des alertes de santé, calculée en fonction de la sévérité des boiteries, des mammites ou pour l'ensemble des métrites est présentée dans le tableau 3-III-5.

Tableau 3-III-5 : sensibilité des alertes de santé et intervalle de confiance à 95% ($IC_{95\%}$) en fonction de la sévérité du trouble clinique.

Classes de boiteries	Nombre	Sensibilité (%)	$IC_{95\%}$
Boiterie modérée	15	13	[0 ; 60]
Boiterie sévère	4	50	[7 ; 93]
Ensemble des boiteries	19	11	[1 ; 33]
Mammite légère	6	33	[4 ; 78]
Mammite modérée	8	25	[3 ; 65]
Mammite sévère	3	67	[9 ; 99]
Ensemble des mammites	17	35	[14 ; 62]
Métrites	13	46	[19 ; 75]
Ensemble des maladies	56	32	[20-46]

La sensibilité des alertes de santé pour la détection des boiteries, mammites et métrites était respectivement de 11, 35 et 46%. Sauf pour les mammites sévères, pour lesquelles la sensibilité

des alertes était plus élevée (67%), la sévérité des boiteries et des mammites n'impactait pas de manière nette la sensibilité des détections. Cependant, les effectifs à l'intérieur de chaque classe de sévérité sont très limités et par conséquent la sensibilité par classe de sévérité doit être interprétée très prudemment.

Parmi les vaches classées non malades, 19% des vaches ne présentant aucune anomalie, 29% des vaches présentant une boiterie légère et 25% des vaches en acétonémie subclinique déclenchaient des alertes de santé.

Parmi les vaches déclenchant des alertes optimisées et sans anomalie signalée par l'éleveur, 12/48 (25%) étaient « malades » (6 métrites, 1 boiterie, 1 hyperthermie, 1 suspicion de péritonite par corps étranger, 1 déplacement de caillette à gauche, 1 trouble digestif et 1 trouble respiratoire).

5. Discussion

La méthode de référence utilisée dans cette étude pour calculer la valeur informative des alertes de santé, l'ECV, a été très peu utilisée pour évaluer un outil de monitoring susceptible de détecter plusieurs types de troubles de santé, et à notre connaissance jamais utilisé chez les vaches laitières. Notre hypothèse était que l'utilisation de cette méthode pouvait permettre d'estimer plus correctement la spécificité des alertes (spécificité attendue plus élevée), certaines alertes pouvant être considérées à tort comme des faux positifs lorsque l'éleveur est utilisé comme méthode de référence, en raison de sa sous-détection de certains troubles (Fabian et al., 2014). Or la spécificité obtenue dans cette étude (79%, $IC_{95\%} = (72-84)$) est inférieure à celle obtenue avec les mêmes algorithmes dans le chapitre 3.I (90%, $IC_{95\%} = (88-92)$), ce qui est contraire au résultat attendu. Plusieurs hypothèses peuvent être formulées pour expliquer ce résultat. Premièrement, la différence observée peut provenir de la variabilité des performances des algorithmes entre fermes (2 fermes pour le chapitre 3.I, une autre ferme utilisée pour le chapitre 3.III), en raison des différences dans les conditions d'élevage (management des vaches, alimentation, accès au pâturage) qui peuvent expliquer certains faux positifs. Ensuite, même si l'utilisation de l'ECV permet de mieux classer certains événements qui ne sont plus considérés à tort comme des faux positifs (par exemple alertes de santé chez des vaches atteintes de métrites non vues par l'éleveur), les faux positifs liés à une mauvaise classification des animaux malades par l'éleveur sont probablement beaucoup moins nombreux que les faux positifs liés à des

facteurs environnementaux impactant la rumination et l'activité (alimentation, accès au pâturage...). Les quelques évènements reclassés correctement en vrai positifs grâce à l'ECV auraient finalement peu d'impact sur le calcul de la spécificité. Enfin, la moins bonne spécificité des alertes de santé optimisées dans cette étude peut s'expliquer par le fait que les paramètres des algorithmes CUSUM, et notamment la valeur de la constante k , ont été optimisés en utilisant les données du chapitre 3.I. Une moins bonne performance des algorithmes pouvait donc être attendue lors de l'application sur un jeu de données indépendant du jeu de données de développement des algorithmes.

Même si aucune hypothèse sur l'effet du changement de méthode de référence sur la sensibilité des alertes ne pouvait être formulée, la sensibilité tous troubles de santé calculée dans cette étude (32%, $IC_{95\%} = (21-45)$) est du même ordre de grandeur que la sensibilité calculée dans le chapitre 3.I (40%, $IC_{95\%} = (35-45)$). Nous pouvons supposer que la détection des maladies par ECV était plus sensible que celle faite par l'éleveur par observation visuelle. En effet, l'ECV reposait sur l'analyse de données physiques et physiologiques relevées sur l'animal après un examen spécialisé systématique de tous les appareils tandis que l'éleveur s'appuie sur des données de performances (production de lait, ingestion) et sur l'observation du comportement de l'animal pour diagnostiquer les maladies. L'ECV était plus systématique (prise de température sur tous les animaux, recherche de métrites à l'aide d'un vaginoscope...). Le tableau 3-III-4 confirme que l'ECV est plus sensible que l'éleveur pour la détection des maladies, avec seulement 37% des maladies qui avaient donné lieu à un signalement d'anomalie par l'éleveur. La sous-détection est particulièrement importante pour les boiteries, avec seulement 5% des boiteries pour lesquelles une anomalie avait été signalée par l'éleveur. Ceci est concordant avec les résultats d'autres études qui ont montré que la prévalence des boiteries dans leur troupeau était en général minorée par les éleveurs, comparé à une notation standardisée par un examinateur entraîné (Fabian et al., 2014, Leach et al., 2010). Pour seulement 15% des métrites, l'éleveur avait signalé une anomalie chez l'animal avant son ECV, ce qui pourrait s'expliquer par la discrétion des signes cliniques associés aux endométrites si l'on n'examine pas l'appareil génital. L'ECV permet également probablement d'aboutir à un diagnostic plus précis que le diagnostic fait par l'éleveur dans un certain nombre de cas (par exemple déplacements de caillette rarement diagnostiqués par les éleveurs). De plus, l'ECV est plus informatif que le repérage des chutes de production laitière (chapitre 3.II) pour déterminer l'état de santé des animaux, puisqu'il aboutit généralement en cas d'anomalie à un diagnostic de maladie alors que l'observation d'une chute de production laitière seule ne permet pas de poser un diagnostic.

Les valeurs informatives obtenues dans cette étude, comparées aux valeurs obtenues dans le chapitre 3.I, indiquent cependant qu'il n'y a pas d'intérêt majeur à réaliser des ECV pour évaluer la sensibilité et la spécificité d'alertes de santé générées par un système de monitoring. En effet, l'ECV, s'il semble être la meilleure méthode de référence pour le diagnostic des troubles cliniques, nécessite de faire participer des vétérinaires au protocole ce qui peut être coûteux. De plus, le nombre d'animaux qui peuvent être examinés par un ou quelques vétérinaires est limité.

Le design de notre étude était également original pour une autre raison. Nous avons choisi de générer dans un premier temps des alertes de santé sensibles, en détectant sur l'ensemble des vaches de l'élevage des chutes de rumination ou d'activité. Dans un second temps, un ECV était effectué sur l'ensemble des vaches en alerte de santé et sur certaines vaches non en alerte, tirées au sort parmi les vaches non en alerte. Notre hypothèse de travail était que les vaches malades présentaient plus de risque de déclencher de manière précoce des alertes de santé que les vaches non malades. Nous avons en effet montré dans une étude précédente (chapitre 3.I) que 50% des chutes d'activité et de rumination détectées par des CUSUM identiques à ceux employés ici avaient lieu avant la détection d'un trouble de santé par les éleveurs, et 30% avaient lieu le même jour. Les vaches en alerte étant examinées jusqu'à 3 jours après le début de l'alerte de santé, ce délai permettait d'examiner les vaches jusqu'à l'apparition des signes cliniques. Cependant, pour certaines alertes de santé de courte durée, il est possible que des vaches aient été examinées une seule fois et cela avant l'apparition des signes cliniques, induisant une mauvaise classification des ECV (faux positifs au lieu de vrais positifs). Le design de cette étude permettait théoriquement de maximiser les chances d'examiner des animaux malades tout en limitant le nombre d'animaux à examiner, un ECV quotidien de l'ensemble des animaux de l'élevage n'étant pas possible pour une question de temps et de moyens. Cependant, en réalité les animaux en alerte de santé ne présentaient pas beaucoup plus de risque d'être malades, puisqu'avec les seuils de CUSUM appliqués initialement, 32% (48/150) des vaches en alerte de santé présentaient une maladie contre 25% (32/126) des vaches non en alerte de santé. Ce faible différentiel ainsi que plus généralement la faible incidence de troubles de santé dans l'élevage inclus dans cette étude, ont fait que peu d'animaux ont été diagnostiqués comme malades (56/240). Ceci implique que les estimations de la sensibilité des alertes de santé tous troubles confondus, et de la sensibilité par trouble pour les principaux troubles de santé sont assez imprécises. L'incidence des troubles de santé ayant des répercussions cliniques en élevage bovin laitier étant généralement modérée (Fourichon et al., 2001), estimer avec une bonne précision la sensibilité d'alertes de santé dans ce type d'élevage est un challenge, d'autant plus que les élevages équipés de systèmes de monitoring sont probablement plutôt des élevages performants

techniquement, où la situation sanitaire n'est pas trop dégradée. Un suivi de davantage d'élevages pendant plus longtemps aurait pu permettre d'améliorer la précision des estimations. Le design d'étude utilisé ici pourrait cependant être utilisé pour estimer la sensibilité d'alertes de santé dans des systèmes d'élevage où l'incidence des troubles de santé est plus importante, par exemple chez des jeunes bovins après l'allotement, en ateliers de veaux de boucherie, ou chez des vaches laitières lorsque la situation sanitaire du troupeau est dégradée (suite à une fusion de troupeaux par exemple)...Le design de l'étude permettait par ailleurs d'estimer la spécificité des détections avec une bonne précision (79% $IC_{95\%} = (72-84)$), alors que la spécificité des systèmes d'alertes n'est pas systématiquement évaluée dans les publications (Huybrechts et al., 2014). Une des limites de cette méthode est que les animaux étant recrutés sur la base des alertes de santé, les troubles n'impactant pas l'activité et la rumination des animaux risquent d'être sous représentés parmi les vaches malades à l'issue de l'ECV.

Les sensibilités par type de trouble de santé sont faibles, variant entre 11% pour les boiteries et 46% pour les métrites. Ce résultat s'explique en partie car les algorithmes CUSUM évalués avaient été paramétrés pour obtenir une spécificité suffisante (chapitre 3.I), en limitant ainsi le nombre d'alertes sur des animaux non malades pour l'éleveur. Un paramétrage des algorithmes de manière à obtenir une sensibilité plus élevée et une spécificité plus faible est possible. Cependant, de nombreuses alertes à tort sont alors générées et la valeur prédictive positive des alertes diminue rapidement (chapitre 3.I). Il n'était par ailleurs pas attendu que tous les troubles de santé provoquent des chutes d'activité ou des durées de rumination. En effet, ces grandeurs sont liées à l'état général des animaux. Il était donc attendu *a priori* qu'elles soient altérées principalement pour les maladies sévères affectant l'état général des animaux. On remarque par exemple que 2 des 3 mammites sévères ont donné lieu à une alerte de santé. Van Hertem et al. (2013) ont évalué les performances de détection d'un modèle basé sur des données d'activité et de rumination issues des capteurs utilisés dans Heatime-HR® pour la détection des boiteries. La sensibilité (89%) et la spécificité (85%) des alertes générées par le modèle étaient élevées. Cependant, des variables relatives à la production laitière des animaux étaient également incluses dans le modèle, ce qui peut en partie expliquer les meilleurs résultats obtenus, les données liées à la production laitière étant plus liées aux boiteries que les données d'activité et de rumination dans cette étude. Par ailleurs, les variables d'activité et de rumination incluses dans le modèle n'étaient pas les données brutes issues du capteur mais des variables créées à partir de ces données brutes. Enfin, des fenêtres très larges étaient utilisées autour des épisodes de boiteries ou des jours témoins (3 semaines avant, 3 semaines après) pour calculer la sensibilité et la spécificité des détections. La précocité des alertes n'étant pas évaluée dans l'étude, il est possible

que certaines alertes générées par le modèle soient à relier au traitement des boiteries (parage agressif avec mise à nu de lésion par exemple) et non à la boiterie elle-même. On peut supposer que certaines boiteries diagnostiquées lors des ECV dans notre étude avaient en réalité commencé plusieurs jours avant l'examen. Une fenêtre plus longue que 4 jours pour la validation des alertes liées aux boiteries aurait donc pu être pertinente. Par ailleurs, certaines boiteries chroniques peuvent probablement engendrer des baisses d'activité générale et des durées de rumination qui sont trop progressives pour être détectées par les algorithmes que nous avons utilisés dans cette étude. La sensibilité des alertes pour la détection des mammites est faible (35%, $IC_{95\%} = (14-62)$). Ceci est conforme au résultat obtenu dans une étude précédente (chapitre 3.I). Ici, la détection des mammites était effectuée par l'éleveur par palpation de la mamelle et examen visuel des premiers jets en cas de suspicion de mammite. L'examen des premiers jets n'ayant pas été systématique, on peut penser que la détection des mammites n'était pas très sensible. La détection des mammites subcliniques par le système d'alertes n'a pas été étudiée dans cette étude car un produit de trempage étant appliqué avant tout examen clinique, il n'était pas possible d'effectuer par exemple un California Mastitis Test lors de l'ECV. Cependant, au vu des résultats obtenus sur les mammites cliniques détectées par l'éleveur, l'hypothèse que les mammites subcliniques puissent déclencher des alertes est peu probable. La sensibilité des alertes dans la détection des métrites est meilleure puisqu'elle est de 46% ($IC_{95\%} = (19-75)$). Ce résultat est intéressant si l'on considère que les métrites sont difficiles à détecter par les éleveurs car les signes cliniques associés sont frustes. De plus l'impact économique des métrites est important, en raison de la dégradation des performances de reproduction engendrée (baisse de fertilité et de fécondité) et d'un risque de réforme accru notamment (Fourichon et al., 2000, LeBlanc et al., 2002). Des études complémentaires seraient nécessaires pour mieux décrire le lien entre chutes d'activité et de rumination et présence de métrites chez les vaches.

Nous avons choisi de classer les vaches présentant uniquement une acétonémie subclinique comme non malades dans cette étude. En effet, le pourcentage d'alertes de santé chez ces vaches n'était pas significativement différent de celui observé chez les vaches ne présentant aucune anomalie à l'ECV et chez les vaches présentant uniquement une boiterie légère. De plus, un modèle d'analyse multivarié construit avec les données issues de cette étude a montré que la mise en évidence d'une acétonémie subclinique chez les vaches n'affectait pas significativement les durées de rumination mesurées (résultats non montrés).

Les sensibilités obtenues dans notre étude pour les alertes de santé sont trop faibles pour considérer que le système d'alertes peut se substituer à la vigilance de l'éleveur. Il permet cependant d'apporter des informations complémentaires à l'éleveur sur la santé de son troupeau.

Dans notre étude, 12 alertes de santé ont ainsi été déclenchées pour des vaches malades non détectées par l'éleveur. Le système d'alertes peut donc permettre d'accroître la surveillance de la santé du troupeau par l'éleveur, en repérant des animaux ayant échappé à la vigilance de celui-ci.

6. Conclusion

Le design original de l'étude n'a pas complètement permis de résoudre une difficulté qui était attendue, à savoir le recrutement de suffisamment d'animaux malades pour calculer la sensibilité des alertes de santé avec une bonne précision. Pour une spécificité suffisante (79%), limitant le nombre de fausses alertes pour l'éleveur, 32% de l'ensemble des troubles de santé ont été détectés par les algorithmes CUSUM utilisés. Seules les métrites ont été détectées avec une sensibilité moyenne (46%). Le design utilisé a cependant permis de confirmer que certaines maladies comme les boiteries ou les métrites étaient sous-diagnostiquées par l'éleveur.

Au vu des contraintes imposées par la réalisation d'ECV systématiques sur les vaches, nos résultats montrent que l'utilisation de cette méthode de référence n'est pas nécessaire pour estimer avec une bonne précision les valeurs informatives d'alertes de santé générées par un système de monitoring.

Références bibliographiques

- Fabian, J., R. A. Laven, and H. R. Whay. 2014. The prevalence of lameness on new zealand dairy farms: A comparison of farmer estimate and locomotion scoring. *Vet J* 201:31-38.
- Fourichon, C., H. Seegers, and X. Malher. 2000. Effect of disease on reproduction in the dairy cow: A meta-analysis. *Theriogenology* 53:1729-1759.
- Fourichon, C., F. Beaudeau, N. Bareille, and H. Seegers. 2001. Incidence of health disorders in dairy farming systems in western france. *Livest Prod Sci* 68:157-170.
- Gernand, E., P. Rehbein, U. U. von Borstel, and S. König. 2012. Incidences of and genetic parameters for mastitis, claw disorders, and common health traits recorded in dairy cattle contract herds. *J Dairy Sci* 95:2144-2156.
- Gillund, P., O. Reksen, Y. T. Gröhn, and K. Karlberg. 2001. Body condition related to ketosis and reproductive performance in norwegian dairy cows. *J Dairy Sci* 84:1390-1396.
- Huybrechts, T., K. Mertens, J. De Baerdemaeker, B. De Ketelaere, and W. Saeys. 2014. Early warnings from automatic milk yield monitoring with online synergistic control. *J Dairy Sci* 97:3371-3381.
- Iwersen, M., U. Falkenberg, R. Voigtsberger, D. Forderung, and W. Heuwieser. 2009. Evaluation of an electronic cowside test to detect subclinical ketosis in dairy cows. *J Dairy Sci* 92:2618-2624.
- Leach, K. A., H. R. Whay, C. M. Maggs, Z. E. Barker, E. S. Paul, A. K. Bell, and D. C. Main. 2010. Working towards a reduction in cattle lameness: 1. Understanding barriers to lameness control on dairy farms. *Res Vet Sci* 89:311-317.
- LeBlanc, S. J., T. F. Duffield, K. E. Leslie, K. G. Bateman, G. P. Keefe, J. S. Walton, and W. H. Johnson. 2002. Defining and diagnosing postpartum clinical endometritis and its impact on reproductive performance in dairy cows. *J Dairy Sci* 85:2223-2236.
- Luo, Y., Z. Li, and Z. Wang. 2009. Adaptive CUSUM control chart with variable sampling intervals. *Comput. Stat. Data Anal.* 53:2693-2701.
- Radostits, O. M., C. C. Gay, K. W. Hinchcliff, and P. D. Constable. 2007. *Veterinary Medicine: A textbook of the diseases of cattle, horses, sheep, pigs, and goats*. 10th ed. Saunders Elsevier, New York

- Reith, S. and S. Hoy. 2012. Relationship between daily rumination time and estrus of dairy cows. *J Dairy Sci* 95:6416-6420.
- Rutten, C. J., A. G. Velthuis, W. Steeneveld, and H. Hogeveen. 2013. Invited review: Sensors to support health management on dairy farms. *J Dairy Sci* 96:1928-1952.
- Schirmann, K., M. A. von Keyserlingk, D. M. Weary, D. M. Veira, and W. Heuwieser. 2009. Technical note: Validation of a system for monitoring rumination in dairy cows. *J Dairy Sci* 92:6052-6055.
- Schirmann, K., N. Chapinal, D. M. Weary, L. Vickers, and M. A. von Keyserlingk. 2013. Short communication: Rumination and feeding behavior before and after calving in dairy cows. *J Dairy Sci* 96 :7088-7092.
- Sepulveda-Varas, P., D. M. Weary, and M. A. G. von Keyserlingk. 2014. Lying behavior and postpartum health status in grazing dairy cows. *J Dairy Sci*. [http://dx.doi.org/ 10.3168/jds.2014-8357](http://dx.doi.org/10.3168/jds.2014-8357).
- Sheldon, I. M., G. S. Lewis, S. LeBlanc, and R. O. Gilbert. 2006. Defining postpartum uterine disease in cattle. *Theriogenology* 65:1516-1530.
- Sprecher, D. J., D. E. Hostetler, and J. B. Kaneene. 1997. A lameness scoring system that uses posture and gait to predict dairy cattle reproductive performance. *Theriogenology* 47:1179-1187.
- Steen, A. 2001. Field study of dairy cows with reduced appetite in early lactation: Clinical examinations, blood and rumen fluid analyses. *Acta Vet Scand* 42:219-228.
- Timsit, E., N. Bareille, H. Seegers, A. Lehebel, and S. Assie. 2011a. Visually undetected fever episodes in newly received beef bulls at a fattening operation: Occurrence, duration, and impact on performance. *J Anim Sci* 89:4272-4280.
- Timsit, E., S. Assie, R. Quiniou, H. Seegers, and N. Bareille. 2011b. Early detection of bovine respiratory disease in young bulls using reticulo-rumen temperature boluses. *Vet J* 190:136-142.
- Van Hertem, T., E. Maltz, A. Antler, C. E. Romanini, S. Viazzi, C. Bahr, A. Schlageter-Tello, C. Lokhorst, D. Berckmans, and I. Halachmi. 2013. Lameness detection based on multivariate continuous sensing of milk yield, rumination, and neck activity. *J Dairy Sci* 96:4286-4298.

Annexes du chapitre 3-III

Annexe 3-III-1

Examen clinique vétérinaire-Gaec du Ranch-Hiver 2014

Animal		
N° travail :	Date ECV :	Heure ECV :
Catégorie (A+A-/E+E-) :	Date vêlage :	Position Tag :

Etat général-Fréquences				
NEC	AR :	TR	FC	FR
	Flanc :	:°C	:/15s	: /min
Vaginoscope (>15j p.p.-IA1)	non <input type="checkbox"/>	0 <input type="checkbox"/>	1 <input type="checkbox"/>	2 <input type="checkbox"/>
		3 <input type="checkbox"/>		
		Hypomotilité <input type="checkbox"/>	Remarques EG :	
Contractions ruminales :		Normales <input type="checkbox"/>		
		Hypermotilité <input type="checkbox"/>		
Muqueuses :	normales	Pâles <input type="checkbox"/>	Congestionnées <input type="checkbox"/>	Ictériques <input type="checkbox"/>
<input type="checkbox"/>				
Froides/Sèches <input type="checkbox"/>		Si modification,		
		intensité +TRC		

Mamelle/Lait				
Mammite	0 <input type="checkbox"/>	1 <input type="checkbox"/>	2 <input type="checkbox"/>	3 <input type="checkbox"/>
Autres :				
Rq : Examiner 1 E+ mammite de tps en tps (post-trempage)				

Appareil cardio-vasculaire				
Œdème	Absence <input type="checkbox"/>	Présence <input type="checkbox"/>	localisation	
Jugulaires				
Auscultation				
Intensité des bruits normaux	normale <input type="checkbox"/>	augmentée <input type="checkbox"/>	diminuée <input type="checkbox"/>	
Bruits surajoutés	liquidiens <input type="checkbox"/>	frottements <input type="checkbox"/>	galop <input type="checkbox"/>	souffle <input type="checkbox"/>
<u>Localisation</u> (EIC+hauteur) :				
Autres :				

Appareil digestif				
Mastication /Déglutition	normale <input type="checkbox"/>	anormale <input type="checkbox"/>	Préciser :	
Succussion	Absence de bruits <input type="checkbox"/>	de Bruits liquidiens <input type="checkbox"/>	Signe flot <input type="checkbox"/>	du Localisation
Ping <input type="checkbox"/>	Localisation			
Contenu rumen	normal <input type="checkbox"/>	Sec <input type="checkbox"/>	gazeux <input type="checkbox"/>	liquide <input type="checkbox"/>
Consistance bouses	sec <input type="checkbox"/>	normal <input type="checkbox"/>	liquide <input type="checkbox"/>	
Couleur bouses	jaune <input type="checkbox"/>	Brun-vert <input type="checkbox"/>	noir <input type="checkbox"/>	
Éléments solides	maïs <input type="checkbox"/>	Fibres- <input type="checkbox"/>	Normal <input type="checkbox"/>	Fibres+ <input type="checkbox"/>
Autres :				

Appareil respiratoire			
Dyspnée	Absence <input type="checkbox"/>	Présence <input type="checkbox"/>	Orthopnée <input type="checkbox"/>
Respiration	Costo-abdominale <input type="checkbox"/>	abdominale <input type="checkbox"/>	thoracique <input type="checkbox"/>
Bruits à distance	cornage <input type="checkbox"/>	toux <input type="checkbox"/>	Préciser
Jetage <input type="checkbox"/>	Décrire		
Auscultation	normale <input type="checkbox"/>	Absence de bruits <input type="checkbox"/>	Bruits normaux augmentés <input type="checkbox"/>
Bruits surajoutés	sifflements <input type="checkbox"/>	crépitements <input type="checkbox"/>	Frictions pleurales <input type="checkbox"/>
Autres :			

Examens complémentaires				
Test Glutaraldéhydemin			
BHB :	mmol/L			
Glycémie :	mg/dL			
Bandelette	oui <input type="checkbox"/>	Non <input type="checkbox"/>	Résultat	
Prélèvements conservatoires	Tube sec <input type="checkbox"/>	He-Li <input type="checkbox"/>		
Sore boiterie	0 <input type="checkbox"/>	1 <input type="checkbox"/>	2 <input type="checkbox"/>	3 <input type="checkbox"/>

Autres

Annexe 3-III-2

Examens complémentaires-Grilles d'examens cliniques

Tableau 3-III-6 : Grille de notation des mérites en fonction des observations au vaginoscope

Score (interprétation)	Critères
0 (absence de mérite)	Aspect du mucus vaginal clair
1 (absence de mérite)	Aspect du mucus trouble avec quelques mèches de pus
2 (mérite)	Décharge de matériel muco-purulent avec moins de 50% de pus
3 (mérite)	Décharge de matériel muco-purulent, blanc, jaune ou sanguinolent avec plus de 50% de pus

Tableau 3-III-7 : Grille de notation des mammites

Score (interprétation)	Critères
1 (légère)	Présence de quelques grumeaux dans le lait
2 (modérée)	Présence de grumeaux dans le lait associé à une induration du (des) quartier(s) atteint(s)
3 (sévère)	Modification de l'aspect du lait associé à une atteinte de l'état général : hyperthermie, baisse d'appétit ou chute de production de lait

Tableau 3-III-8 : Grille de notation des boiteries

Score (interprétation)	Démarche irrégulière	Ligne du dos arquée en mouvement	Ligne du dos arquée à l'arrêt	Enjambées courtes	Mouvements de la tête	Membre affecté identifiable	Réticence à supporter son poids
0 (non boiteuse)	non	non	non	non	non	non	non
1 (légère)	oui	oui	+/-	+/-	non	non	non
2 (modérée)	oui	oui	oui	oui	oui	oui	non
3 (sévère)	oui	oui	oui	oui	oui	oui	oui

Chapitre 4. Discussion générale

De nombreux outils de monitoring de la reproduction et de la santé ont été développés ces dernières années pour aider les éleveurs bovins pour la détection d'évènements physiologiques (chaleurs, vêlages) ou pathologiques (mammites, boiteries). Le développement de ces outils coûte cher donc ils sont souvent mis sur le marché avant que leurs potentialités et leurs performances aient été complètement évaluées. Cette thèse était basée sur l'utilisation de Heatime-HR®, un outil de monitoring mesurant en continu l'activité générale et les durées de rumination des vaches laitières. Pour l'entreprise de sélection bovine finançant cette thèse, il était important d'évaluer si les durées de rumination ainsi mesurées pouvaient permettre d'estimer indirectement les quantités ingérées et/ou l'efficacité alimentaire des vaches laitières. Les performances de l'outil pour la détection des troubles de santé étaient également importantes à connaître pour accompagner sa mise sur le marché.

Notre problématique de recherche était de déterminer quelle était le potentiel des grandeurs mesurées par l'appareil Heatime-HR® (activité générale et durées de rumination), ou d'alertes générées à partir de ces grandeurs, pour la détection des troubles de santé ou le phénotypage.

Les principaux objectifs de cette thèse étaient 1) d'évaluer l'intérêt des durées de rumination fournies par Heatime-HR® pour estimer les quantités d'aliment ingérées par des vaches en lactation 2) d'évaluer la valeur informative des données élémentaires de rumination et d'activité fournies par l'appareil Heatime-HR® pour la détection des troubles de santé. Pour cela, des algorithmes spécifiques de détection des troubles de santé, différents des algorithmes commerciaux, ont été développés et évalués en utilisant comme méthode de référence successivement la détection des troubles par observation visuelle par les éleveurs, la détection de chutes importantes de production laitière et enfin un examen clinique vétérinaire systématique de vaches présentant ou non des chutes de rumination ou d'activité.

La première partie de ce chapitre résume les principaux résultats de cette thèse. Les intérêts et les limites des approches retenues sont discutées dans une seconde partie. Enfin, les perspectives ouvertes par ce travail de thèse sont abordées dans une troisième partie.

1. Principaux résultats

1.1 La connaissance des durées de rumination ne permet pas de mieux prédire les quantités ingérées

Le premier résultat majeur de cette thèse a été de montrer que **la connaissance des durées de rumination ne permettait pas de mieux prédire les quantités ingérées par des vaches laitières** en lactation. En effet, lors d'un essai mené sur 41 vaches en lactation pendant 10 semaines (**chapitre 2**), nous avons observé qu'il n'existait pas de relation entre les quantités de matière sèche moyennes ingérées par les vaches et les durées de rumination mesurées par l'outil Heatime-HR®. Certaines vaches avaient ingéré peu de matière sèche en moyenne au cours de l'essai et ruminé beaucoup et inversement. Nous avons également montré que même si les quantités ingérées augmentaient de 0,031 kg par heure de rumination, cet effet était trop faible pour augmenter de manière importante la qualité de prédiction d'un modèle couramment utilisé pour prédire les quantités ingérées (NRC, 2001). Même si l'échantillon d'étude était de taille limitée, les résultats du **chapitre 2** semblent suffisamment clairs pour conclure à l'absence d'intérêt de la mesure des durées de rumination pour estimer les quantités ingérées au cours de la lactation. En effet, l'essai en question s'est déroulé dans des conditions très favorables pour mettre en évidence un lien entre durées de rumination et quantités ingérées puisque les vaches recevaient toutes une ration identique et stable durant tout l'essai. En conditions d'élevage, la ration n'est pas forcément aussi stable (part d'herbe dans la ration par exemple), donc la relation entre quantités ingérées et durées de rumination pourrait être encore moins forte. Par ailleurs, nos résultats sont concordants avec ceux de Schirmann et al. (2012) qui n'avaient pas trouvé de corrélation entre quantités ingérées et durées de rumination au cours du tarissement.

Les résultats de ce chapitre, et plus généralement l'analyse des données de rumination récoltées au cours de cette thèse, montrent une **variabilité importante des durées de rumination entre vaches** (entre 3,8 et 9,7 h/jour en moyenne sur 10 semaines pour des vaches recevant la même ration dans le chapitre 2). Même si une variabilité un peu moins importante a été trouvée par certains auteurs (Soriani et al., 2012) utilisant la même méthode de mesure, nous avons peu d'explications à la variabilité constatée. Des différences anatomiques (longueur de la mâchoire, taille du rumen ou de l'orifice réticulo-omasal) ou

physiologiques (mécanismes de contrôle de la rumination, fréquence de mastication) pourraient expliquer une partie de cette variabilité. Même si la mesure des durées de rumination par Heatime-HR® chez la vache laitière a été jugée fiable dans les quelques études menées (cf chapitre 1), en conditions de terrain (mesure en continu durant de longues périodes) il est possible que l'erreur commise lors de la mesure soit plus importante sur certains animaux (animaux pour lesquels les sons liés à la régurgitation ou les fréquences de régurgitations s'écartent des paramètres utilisés pour régler l'outil par exemple). Une partie de la variabilité des durées de rumination pourrait donc être liée à la technique de mesure et non à de la variabilité biologique. **A l'échelle de la journée et intra-vache, la variabilité des durées de rumination est plus réduite**, même si le coefficient de variation obtenu pour les durées de rumination est 2 fois plus important que celui obtenu pour la production laitière ou les quantités ingérées (chapitre 2). En accord avec les résultats d'autres études (Adin et al., 2009, Schirmann et al., 2012, Soriani et al., 2012) **la variabilité des durées de rumination sur le nyctémère observée dans nos jeux de données est importante**, la rumination s'effectuant par exemple majoritairement la nuit. Concernant **l'activité générale mesurée par Heatime-HR®, elle apparaît également extrêmement variable entre vaches, et présente intra-vache des variations nyctémérales assez répétables**.

Ainsi, le chapitre 2 et les analyses de données réalisées préalablement aux différents chapitres de cette thèse ont montré qu'**entre vaches, la variabilité des durées de rumination et de l'activité générale était importante. Intra-vache, ces 2 grandeurs présentent des variations nyctémérales assez répétables**. C'est pourquoi nous avons fait le choix de prédire les valeurs d'activité et de rumination attendues pour une vache donnée à partir des valeurs observées dans les jours précédents et à la même heure de la journée chez cette même vache.

1.2 Les alertes de santé développées ne se substituent pas à l'observation des troubles de santé par l'éleveur

Le deuxième résultat de cette thèse a été de montrer que des alertes de santé (développées au cours de cette thèse, différentes des alertes commerciales) générées à partir des données d'activité et de rumination mesurées en continu sur les vaches laitières étaient complémentaires de la détection par observation visuelle par les éleveurs pour la détection des

troubles de santé, et ne pouvaient en aucun cas se substituer à celle-ci. En effet, pour une spécificité acceptable pour une utilisation quotidienne, la sensibilité des alertes de santé par rapport à la détection par l'éleveur était moyenne (40% de la totalité des troubles détectés) (**chapitre 3.I**). Pour les troubles de santé sans impact sur l'état général des vaches, les sensibilités trouvées étaient faibles à moyennes (<40%). Pour les troubles de santé ayant un impact sur l'état général (mammites sévères, autres troubles hyperthermisants), les sensibilités trouvées étaient généralement supérieures à 60%. Globalement, pour de tels niveaux de sensibilité, l'éleveur ne peut pas se reposer entièrement sur ces alertes de santé pour la détection des troubles de santé, contrairement à ce qui est observé pour la détection des chaleurs par exemple, où les performances obtenues par les systèmes automatisés sont équivalentes voire supérieures aux performances de détection des éleveurs (At-Taras et Spahr, 2001, Rutten et al., 2013a). La sensibilité moyenne des alertes de santé peut s'expliquer par plusieurs raisons. Tout d'abord, contrairement à un grand nombre de travaux menés sur l'évaluation de systèmes automatisés des troubles de santé, nous avons choisi de paramétrer nos algorithmes de manière à obtenir une spécificité élevée (cf chapitre 4 -3.2) ce qui diminue mécaniquement la sensibilité des alertes. Certains modèles de détection peuvent présenter une sensibilité acceptable, mais avec une spécificité associée très faible. Ainsi l'algorithme utilisé par Adams et al. (2013) détectait 67% des mammites de l'échantillon, mais avec une spécificité associée de seulement 77%. A ces niveaux de spécificité, de nombreux faux positifs sont générés et l'applicabilité des algorithmes en élevage est questionnable. C'est pourquoi le choix de certains auteurs de paramétrer leurs algorithmes de manière à obtenir une sensibilité minimale élevée (Miekley et al., 2013) ne nous paraît pas pertinent au regard de la faible spécificité généralement obtenue. Nous avons par exemple montré dans le chapitre 3.I qu'il était possible d'obtenir une sensibilité élevée des alertes de santé (réglage 3 des CUSUM : respectivement 74% des mammites, 74% des boiteries, 82% des atteintes de l'état général sans cause identifiée détectées), mais avec une spécificité associée insuffisante (spécificité = 51%, 19 alertes/jour pour 100 vaches présentes en l'absence de trouble de santé). Ensuite, les 2 grandeurs mesurées par l'appareil sont plutôt liées à l'état général des vaches, et l'on s'attend donc à ce qu'elles soient diminuées plutôt pour les maladies sévères, avec atteinte de l'état général. Or nous avons vu que la majorité des troubles observés dans nos études était constituée par les mammites, une majorité d'entre elles ayant des répercussions principalement locales sur la mamelle (chapitre 3.I), associées à des signes cliniques généraux d'intensité modérée. Il n'est donc pas étonnant que ces mammites

« locales » soient plus difficilement détectables par les algorithmes. La sensibilité faible des alertes peut également être expliquée par les nombreux facteurs autres que les troubles de santé pouvant provoquer des chutes d'activité (changement de lot de vache, changement de pâtures associés à des changements de durées de déplacement des vaches...) et des durées de rumination (alimentation (Adin et al., 2009), stress thermique (Soriani et al., 2013), changement de lot (Schirmann et al., 2011)...). Tous ces facteurs n'étaient pas contrôlés dans nos études et pouvaient être à l'origine de faux positifs. Pour disposer d'alertes suffisamment spécifiques, utilisables par l'éleveur (cf chapitre 4 -3.2), il fallait donc des seuils de CUSUM relativement élevés, ce qui explique la faible sensibilité obtenue. La fréquence plus élevée des alertes de santé (baisse d'activité ou de la durée de rumination) le jour de la détection d'un événement par la méthode de référence (détection par observation visuelle par l'éleveur ou détection de chutes majeures de production laitière) ou le jour précédent (chapitre 3.I et 3.II) indique cependant que dans ce cas, les alertes de santé sont bien liées aux dégradations de l'état de santé repérées par la méthode de référence.

Les performances moyennes obtenues par les algorithmes développés au cours de cette thèse doivent être remis dans le contexte de la détection automatisée des troubles de santé. Globalement, **les systèmes d'alertes automatisés de santé atteignent rarement à la fois une sensibilité et une spécificité élevée**. Même pour les systèmes dédiés spécifiquement à la détection d'une maladie comme les mammites ou les boiteries, la sensibilité des alertes est rarement supérieure à 50% lorsque leur spécificité est supérieure à 90% (Rutten et al., 2013a). Ainsi, seuls 4 systèmes d'alerte parmi les 28 cités dans cette étude atteignent les performances recommandées par la norme ISO/FDIS 20966 (ISO, 2007) (sensibilité > 80% et spécificité \geq 99%) pour la détection des mammites par les robots de traite. Pour des systèmes d'alerte susceptibles de détecter différentes maladies, comparables au système d'alerte évalué dans cette thèse, des performances comparables ou un peu meilleures sont généralement obtenues (Lukas et al., 2009, Miekley et al., 2013). Cependant, au cours de cette thèse, notre objectif était d'évaluer spécifiquement les performances d'alertes générées à partir de mesures effectuées par un système de monitoring seul, celui-ci étant vendu généralement comme un système autonome (non rattaché à un robot de traite ou à des compteurs à lait) dans la plupart des élevages en France. Certaines études combinent d'emblée dans les modèles de détection des données issues de capteurs commercialisés par un fabricant, et des données issues d'autres composantes du système d'information de l'élevage comme la production laitière. Ainsi, Van

Hertem et al (2013) ont obtenu de bonnes performances avec un modèle de détection des boiteries combinant les données d'activité et de rumination issues du Heatime-HR® avec des données liées à la production laitière journalière. Cependant dans cette étude, les variables liées à la production laitière étaient plus liées aux boiteries que les données d'activité et de rumination. Ce type de modèle obtient donc forcément de meilleures performances. L'intégration des données de production laitière à notre système d'alerte est une perspective de recherche intéressante, vu les performances obtenues dans d'autres travaux (Lukas et al., 2009, Miekley et al., 2013). Cependant en pratique, à part dans les élevages équipés de robots de traite, l'intégration d'informations provenant de différents capteurs ou systèmes d'information est pour le moment limitée par des contraintes techniques (formats de données, interopérabilité...).

Les valeurs prédictives négatives élevées (>99%) trouvées dans le chapitre 3.I semblent indiquer que l'éleveur pourrait se reposer complètement sur l'outil pour la détection des vaches malades, puisque très peu de vaches non en alerte sont malades. Cependant, ce test à la valeur prédictive un peu imparfaite étant répété dans le temps dans une situation où la majorité des vaches de l'élevage n'est pas en alerte de santé, au final 60% des troubles de santé ne donnent pas lieu à une alerte de santé. En raison de la faible prévalence de troubles de santé et de la spécificité imparfaite des alertes, la valeur prédictive positive d'une alerte de santé était inférieure à 10% dans notre échantillon d'étude. Cela confirme que la détection des troubles de santé à partir des alertes générées nécessiterait d'examiner de nombreuses vaches saines (observation visuelle des vaches en alerte, examen clinique, utilisation d'autres données du système d'information de l'élevage pour décider des vaches à examiner suivant les cas) pour pouvoir détecter les rares vaches malades. Ces examens de vaches saines peuvent représenter une perte de temps importante pour l'éleveur et désorganiser son travail. Nous avons également constaté dans le **chapitre 3.I** que **la valeur prédictive positive des alertes diminuait rapidement lorsque la spécificité des alertes diminuait**, pour un gain en valeur prédictive négative très limité. Ainsi, la valeur prédictive positive diminuait de 37% lorsque la spécificité du CUSUM sur la rumination ou l'activité passait de 97% à 90%, et elle diminuait de 59% lorsque la spécificité du CUSUM sur la rumination ou l'activité passait de 90% à 51%. Ce résultat montre que pour les systèmes développés pour la détection des troubles de santé en élevage de vaches laitières, la sensibilité des alertes doit être maximisée pour une spécificité suffisante. En présence d'alertes insuffisamment spécifiques, compte tenu de la faible incidence des troubles cliniques (Fourichon et al., 2001) en élevage laitier,

l'éleveur recevra beaucoup d'alertes sur des animaux sains, ce qui risque de diminuer la confiance qu'il accorde aux informations fournies par le système d'alerte.

Dans le **chapitre 3.I**, nous avons montré que la spécificité des alertes de santé était plus faible en *peripartum*, en raison des chutes d'activité et de rumination observées physiologiquement autour du vêlage. De nombreux troubles de santé ayant lieu autour du vêlage (Gernand et al., 2012), estimer la spécificité des alertes de santé à ce moment-là est important. Pourtant, certains auteurs se contentent d'évaluer les performances de leurs modèles de détection soit globalement, sans prendre en compte d'éventuelles perturbations spécifiquement liées au vêlage, soit uniquement lorsque la lactation est bien établie (Adams et al., 2013, Van Hertem et al., 2013). Ceci s'explique en partie car la collecte des données de monitoring peut parfois être impossible chez les vaches tarées pour des raisons techniques (vaches tarées au pré en permanence par exemple). Les algorithmes de détection peuvent dans ce cas n'être efficaces pour la détection qu'après une durée de plusieurs jours après le vêlage, un délai étant nécessaire pour l'étalonnage des algorithmes. Pour autant, la moins bonne spécificité que nous avons trouvée en *peripartum* ne signifie pas qu'un paramétrage spécifique des algorithmes (avec des seuils h supérieurs par exemple) doit être proposé en *peripartum*. En effet, les vêlages étant généralement étalés en élevage laitier, une spécificité dégradée en *peripartum* n'entraînera pas un grand nombre d'alertes supplémentaires pour l'éleveur au quotidien. Par contre, au vu de la fréquence et de l'impact (pertes en lait) potentiellement important des troubles de santé à cette période, il est important qu'il examine les vaches en alerte en *peripartum*, même si une plus forte proportion de vaches sera non malade.

Au vu des performances moyennes atteintes par les algorithmes développés dans cette thèse, il semble que ce soit la détection des troubles de santé par les alertes fournies par le système d'alerte, **en complément de l'observation visuelle par l'éleveur**, qui soit la méthode optimale de détection. Un résultat identique a été rapporté ailleurs pour la détection des chaleurs (Holman et al., 2011).

1.3 Des alertes de santé plutôt précoces

Les résultats obtenus dans 2 de nos études (**chapitres 3.I et 3.II**) sont concordants et indiquent que lorsque des troubles de santé ou des chutes de production laitière sont détectées par les algorithmes sur l'activité générale ou les durées de rumination, ces détections ont plutôt lieu de manière précoce par rapport à la détection par les éleveurs (30% le même jour que la détection par les éleveurs, 50% dans les 3 jours précédents) ou par rapport au début des baisses de production laitière. Il semble donc que dans certains cas, il soit possible de générer des alertes de santé précoces à partir des données d'activité et de rumination, et donc d'espérer augmenter les chances de guérison et limiter l'impact économique des troubles détectés en limitant leur expression clinique. Cependant, il faut noter que les effets d'un traitement précoce des maladies sur les chances de guérison ne semblent pas systématiques et dépendent des troubles de santé considérés. Si très peu d'études sur l'effet d'un traitement plus ou moins précoce des maladies sont disponibles chez les bovins, des travaux menés chez la souris (Ferran et al., 2011) ont montré que le traitement précoce (10 h versus 32 h post-infection) de pneumonies expérimentales à *Pasteurella multocida* permettaient de réduire la probabilité d'émergence de résistances bactériennes au niveau du poumon, et de diminuer jusqu'à 40 fois la dose d'antibiotique nécessaire pour obtenir une guérison clinique et bactériologique. Des travaux complémentaires sont en cours au sein de l'UMR Bioepar pour évaluer dans quelle mesure ces résultats sont transposables aux bronchopneumonies infectieuses des jeunes bovins par exemple, la détection précoce (12 à 136 h avant l'apparition des signes cliniques) pouvant dans ce cas être assurée par des dispositifs de mesure automatisée de la température corporelle (Timsit et al., 2011a). Chez les vaches laitières, une étude récente (Sannmann et al., 2013, 2014) a montré que le pourcentage de guérison, la concentration en marqueurs de l'inflammation et la production laitière n'étaient pas significativement différents chez 2 groupes de vaches atteintes de métrites puerpérales, 1 groupe étant traité plus précocement (entre 1 et 4 jours postpartum) que l'autre (entre 5 et 10 jours postpartum). De plus dans cette étude, 16% (10/64) des vaches devant initialement être traitées tardivement n'ont pas reçu de traitement car elles n'ont plus présenté de signes cliniques entre 5 et 10 jours postpartum. Ces vaches ont été considérées par les auteurs comme ayant contrôlé spontanément leur infection (guérison spontanée). Plusieurs études ont également étudié l'effet d'un traitement précoce sur les taux de guérison des mammites. Van den Borne et al. (2010) ont par exemple montré que le taux de guérison associé à un

traitement précoces des mammites à *Staphylococcus aureus* tendait à être supérieur à celui observé lors de traitement tardif. Une autre étude s'est intéressée à l'effet d'un traitement précoce de mammites générées par infusion de *Streptococcus uberis* dans un quartier chez 24 vaches laitières (Hillerton et Semmens, 1999). Les 6 vaches traitées au moment de l'apparition des grumeaux dans le lait ont guéri après un traitement antibiotique intramammaire de 10 jours, alors que parmi les 5 vaches traitées précocément suite à une élévation de la conductivité électrique du lait, aucune n'a développé de mammite clinique et la guérison a eu lieu après 7 jours en moyenne. Un autre travail a également étudié l'effet d'un traitement précoce, sur la base d'une élévation de la conductivité électrique du lait, de mammites à *Streptococcus uberis* et à *Staphylococcus aureus*, générées expérimentalement (Milner et al., 1997). Par rapport à un traitement effectué à l'apparition des signes cliniques, le traitement précoce prévenait l'apparition de signes cliniques, diminuait les pertes en lait, diminuait les concentrations en cellules somatiques et diminuait par 2 le temps de guérison pour les 2 pathogènes. Le volume d'antibiotiques utilisés lors de traitement précoce était aussi divisé par 2. Ces différentes études montrent que l'intérêt d'un traitement précoce est probablement variable suivant le type de maladie considéré. Dans certains cas, comme pour les mammites, un traitement précoce peut limiter les répercussions cliniques et zootechniques des maladies. Cependant, un traitement précoce des animaux risque également d'aboutir à des traitements d'animaux en excès (animaux qui auraient guéri spontanément ou animaux faux positifs par le système de détection précoce, si celui-ci manque de spécificité). La diminution du volume d'antibiotique trouvée par Milner et al. (1997) lors de traitement précoce doit probablement être interprétée avec précaution : en ferme, la spécificité imparfaite des alertes de mammites basée sur la conductivité électrique du lait fait qu'un traitement systématique lors d'alertes de mammite pourrait en réalité engendrer une augmentation du volume d'antibiotiques utilisé à l'échelle du troupeau. Plus généralement, la précocité des alertes conduit à se demander quelle doit être l'attitude de l'éleveur lors d'alerte de santé, les signes cliniques pouvant apparaître 1 à 5 jours après l'alerte. Pour l'éleveur, il convient probablement de surveiller ou d'examiner les vaches en alerte pendant quelques jours en cas d'alerte de santé, afin de surveiller l'apparition de signes cliniques (cf chapitre 4 -4.1).

1.4 Comment évaluer les outils de monitoring : quels évènements veut-on détecter ?

1.4.1 Les questionnements initiaux autour de la méthode de référence

Evaluer un système d'alerte utilisant des mesures continues d'activité et de rumination, dont les potentialités en termes de détection des troubles de santé (nature des troubles détectables, performances par type de trouble) étaient largement inconnues au début de cette thèse nous a posé des difficultés, et ce d'autant plus que de nombreux troubles pouvaient potentiellement impacter ces 2 mesures (Radostits et al., 2007). En particulier plusieurs méthodes de référence pouvaient être utilisées, toutes largement imparfaites, c'est-à-dire classant parfois mal le statut de santé des vaches. Il fallait préalablement à toute mise en place de protocole s'interroger sur :

- **L'utilisation finale souhaitée de l'outil** : alertes générées en continu en situation d'élevage et à destination des éleveurs. Les outils de monitoring fournissent des informations d'aide à la décision pour l'éleveur. Les alertes fournies doivent donc déboucher sur des actions concrètes de l'éleveur. Un système qui détecterait en continu les mammites subcliniques n'a probablement pas d'intérêt pour l'éleveur, sauf pour la gestion de la concentration en cellules somatiques du lait de tank (Kamphuis et al., 2013). Pour la gestion des mammites cliniques, il inciterait l'éleveur à traiter ces mammites subcliniques dont une partie aurait guéri spontanément. Un traitement systématique des mammites subcliniques est par ailleurs contraire à la demande sociétale de réduction des intrants médicamenteux en élevage. Evaluer un outil de monitoring vis-à-vis de la détection en continu des mammites subcliniques n'est donc pas pertinent, car cela reviendrait à paramétrer un système beaucoup trop sensible pour l'éleveur, qui recevrait de nombreuses alertes sans utilité. Dans certains cas, pourtant, utiliser comme méthode de référence une méthode jugée plus sensible que l'éleveur peut être pertinent. Ainsi, une étude menée chez des jeunes bovins à l'engraissement (Timsit et al., 2011b), a montré que le gain moyen quotidien des animaux ayant présenté des hyperthermies détectées par un thermobolus ruminal mais non détectées par les éleveurs était significativement diminué (-33 g/j par jour passé en hyperthermie, pendant les 40 jours suivant l'allotement). Une meilleure méthode de référence pour paramétrer un système d'alerte pourrait donc permettre dans certains cas de limiter les pertes zootechniques associées

à certaines maladies en élevage, en permettant un traitement précoce des animaux à un stade subclinique. Il y a là encore risque de conflit d'intérêt entre les éleveurs qui souhaitent sécuriser leur production dans un contexte de volatilité des prix et de réduction des marges de profit (Bewley, 2013), la demande sociétale de réduction des intrants médicamenteux en élevage, et une demande également sociétale de respect de la santé et du bien-être des animaux d'élevage. D'un autre côté, paramétrer des alertes très sensibles expose également au risque de signaler à l'éleveur des animaux qui auraient guéri de leur maladie sans son intervention (guérison spontanée). Etre alerté au sujet de ces animaux représente une surcharge mentale et physique (manipulation, examen, traitement des animaux en alerte) pour l'éleveur et également l'engagement de frais (produits vétérinaires par exemple) qui auraient pu être évités.

• **Les maladies ou évènements qui devaient être en priorité détectées par l'outil** sur la base de leur fréquence, de leur impact économique (pertes zootechniques, frais de traitements, réforme, mort) par exemple. L'objectif est-il de générer des alertes pouvant « remplacer l'œil de l'éleveur » pour les maladies courantes de production ou pouvant aider à la détection des maladies ou signes cliniques difficilement diagnostiqués par l'éleveur (acétonémie, acidose, métrites, chutes de production laitière...). Les systèmes de monitoring fournissent également de nouvelles données en continu, qui remettent en cause les connaissances de ce qui est normal ou pathologique chez l'animal. Ainsi l'utilisation des thermobolus par exemple, ont montré une variabilité des températures corporelles entre bovins et de variations nyctémérales pour un animal donné. Cette variabilité complique la définition simple d'un évènement pathologique (hyperthermie à un seuil de température identique pour tous les animaux par exemple). De même, cela aboutit à isoler un certain nombre d'anomalies (hyperthermies) sans signes cliniques de maladie associés (Timsit et al., 2011a). Ces évènements peuvent dans certains cas être considérés comme des guérisons spontanées, ne devant pas être signalées systématiquement à l'éleveur.

• **L'intégration de l'outil dans le système d'information de l'éleveur** : pour certaines maladies comme les mammites, la méthode de détection de choix au quotidien est probablement la détection par l'éleveur, les vaches passant dans la grande majorité des cas en salle de traite 2 fois par jour. Des informations issues de systèmes de monitoring ne sont pas forcément pertinentes pour tous les troubles de santé.

• **La fiabilité et le coût des méthodes de référence utilisables pour évaluer les états de santé des vaches laitières.** Il a été montré par exemple que la détection des boiteries par les éleveurs était très peu sensible (Fabian et al., 2014). Cette méthode n'est donc pas à privilégier pour évaluer un système détectant uniquement les boiteries.

Ces questionnements nous ont conduits à nous interroger sur le type d'évènements de santé que l'on souhaitait pouvoir détecter avec cet outil, et donc à utiliser successivement 3 méthodes de référence différentes pour l'évaluation des alertes de santé.

1.4.2 Intérêts et limites des 3 méthodes de référence utilisées

Dans une première étude (chapitre 3.I), nous avons utilisé la **détection par observation visuelle par les éleveurs comme méthode de référence**. Avec les alertes présentant le meilleur compromis entre sensibilité et spécificité, 40% de l'ensemble des troubles étaient détectés, avec des sensibilités supérieures à 60% pour les troubles impactant l'état général des animaux. Les troubles rares mais graves étaient généralement détectés. Les intérêts liés à l'utilisation de cette méthode sont qu'elle est identique à la méthode utilisée classiquement pour la détection des troubles de santé en élevage, qu'elle permet de travailler d'emblée sur plusieurs maladies, et qu'elle est peu coûteuse (par rapport à des analyses de laboratoire répétées dans le temps par exemple), ce qui permet de suivre de nombreux animaux durant une période de temps importante. Cet aspect est important étant donné la faible incidence des troubles cliniques en élevage laitier (Fourichon et al., 2001). Les limites sont qu'elle peut être peu sensible pour la détection de certaines maladies comme les boiteries par exemple (Fabian et al., 2014), et que la qualité de l'observation des animaux peut varier dans le temps, par exemple en fonction du temps que les éleveurs peuvent consacrer à l'observation et au traitement des animaux. La qualité de la détection peut également être variable entre éleveurs, en fonction de leur niveau de compétence et d'expérience, des types de troubles de santé qu'ils rencontrent le plus dans leur élevage, de leur plus grande sensibilité à une maladie particulière ou encore du nombre d'animaux atteints qu'ils tolèrent. Les études portant en général sur une ou quelques exploitations, le niveau de détection des éleveurs sélectionnés doit être suffisant pour permettre une analyse pertinente des données. Les performances vis-à-

vis des maladies difficilement diagnostiquées par l'éleveur (maladies métaboliques par exemple) seront plus difficilement évaluées. Cette méthode est également dépendante de la qualité des enregistrements sanitaires, qui n'est pas systématiquement optimale dans toutes les fermes.

Dans le chapitre 3.II, nous avons utilisé simultanément 2 méthodes de référence : la détection par l'éleveur et **le repérage de chutes majeures de production laitière**. Les valeurs informatives trouvées pour les alertes de santé en comparaison avec la détection éleveur (spécificité = 95 %, sensibilité = 25 %) sont différentes de celles du chapitre 3.I (spécificité = 90%, sensibilité = 40%). La plus forte spécificité trouvée dans le chapitre 3.II peut s'expliquer par le fait que les fenêtres temporelles tirées au sort pour ce calcul étaient situées à distance des troubles de santé détectés par l'éleveur et des chutes de lait (versus uniquement à distance des troubles de santé pour le chapitre 3.I). La plus faible sensibilité peut s'expliquer par le fait que les troubles de santé *peripartum*, mieux détectés, étaient exclus de l'analyse dans le chapitre 3.II (la lactation devait être établie pour le repérage des chutes de production laitière). Parallèlement, 36% des chutes majeures de production laitière donnaient lieu à une alerte de santé. Les avantages de l'utilisation des chutes de production comme méthode de référence sont que cela permet de chercher à détecter en priorité les troubles de santé qui ont un impact économique du fait des pertes de lait engendrées (en moyenne 41 kg par chute de production laitière), que la production laitière journalière est une donnée objective dont la mesure est de plus en plus abordable, et qu'une baisse précoce de production laitière est fréquemment observée en relation avec une large variété de troubles de santé (Bareille et al., 2003, Fourichon et al., 1999). La principale limite de cette méthode de référence est qu'il est difficile de déterminer un seuil au-delà duquel une baisse de production laitière doit être détectée. De plus, cette méthode ne donne pas d'informations sur le type de maladies ou de facteurs environnementaux (stress thermique par exemple) ayant affecté la production laitière.

Dans le chapitre 3.III, nous avons utilisé comme méthode de référence **un examen clinique vétérinaire systématique** pour calculer la valeur informative des alertes de santé. Les valeurs informatives obtenues (spécificité = 79%, sensibilité = 32%) montrent que, dans un contexte de faible incidence des troubles de santé (chez les vaches laitières) et compte tenu du coût engendré pour faire examiner de manière répétée un grand nombre d'animaux par des

vétérinaires, il y a peu d'intérêt à utiliser cette méthode de référence de façon systématique pour l'évaluation d'un système fournissant des alertes de santé. Cette méthode de référence et plus généralement le design d'étude utilisé dans le chapitre 3.III pourraient néanmoins présenter un intérêt dans des contextes où l'incidence des troubles de santé est plus élevée (ateliers d'élevage de jeunes bovins par exemple), ou si l'on souhaite poser un diagnostic le plus précis possible pour les affections rencontrées.

1.4.3 Que conclure de la confrontation des résultats obtenus avec les 3 méthodes de référence utilisées ?

La confrontation des résultats obtenus dans ces 3 chapitres permet de tirer des enseignements généraux pour l'évaluation d'un système de monitoring global de la santé (susceptible de détecter une large variété de troubles de santé) en général et d'un système basé sur la mesure de l'activité générale et des durées de rumination en particulier. Tout d'abord, **la combinaison de plusieurs méthodes de référence** (si possible en parallèle sur le même jeu de données) **peut permettre d'évaluer le plus complètement possible quels événements sont détectables** à l'aide du système. Nous avons montré que des chutes majeures et brutales de production laitière (0,7 chute par lactation) n'étaient associées à aucune détection de trouble de santé par les éleveurs (**chapitre 3.II**). Une partie de ces chutes majeures de production laitière (0,2 chute par lactation) coïncidaient avec des baisses d'activité générale et/ou de durées de rumination. Ce résultat montre qu'un outil de monitoring de l'activité générale et de la rumination peut permettre de détecter des baisses brutales de production laitière, correspondant très probablement à une altération de la santé des vaches, alors que l'éleveur seul n'aurait rien détecté. Un examen clinique vétérinaire approfondi de vaches ayant présenté des chutes de rumination ou d'activité (**chapitre 3.III**) a par ailleurs permis de mettre en évidence que certaines vaches examinées présentaient un trouble clinique alors qu'aucune anomalie n'avait été détectée par l'éleveur. La combinaison de différentes méthodes de référence permet également d'être plus rigoureux pour le calcul des valeurs informatives, et en particulier de la spécificité. Nous avons par exemple trouvé 5 point d'écart pour la spécificité des alertes de santé entre le chapitre 3.I (spécificité = 90%), dans lequel les périodes témoins étaient tirées au sort uniquement à distance des troubles de santé détectés par les éleveurs, et le chapitre 3.II (spécificité = 95%), dans lequel les périodes témoin étaient

tirées au sort à distance des détections par les éleveur et des chutes de production laitière. Pouvoir détecter les baisses de production laitière permet dans ce cas d'avoir des périodes témoins plus « propres », dans lesquelles la probabilité que l'animal soit malade est réduite. En revanche, le fait que les **ordres de grandeur des valeurs informatives trouvées dans les 3 études soient proches** (spécificité = 80-95%, sensibilité = 20-40%) **plaide plutôt en faveur de l'utilisation d'une seule méthode de référence pour évaluer un système d'alerte**. Il s'agit alors de bien définir quels types d'évènements on souhaite détecter. **Des biais de classement sont commis lorsqu'une méthode de référence imparfaite est utilisée**. Dans le chapitre 3.I, compte tenu de la sous-détection de certaines maladies par les éleveurs (Fabian et al., 2014, Timsit et al., 2011b), on peut supposer qu'un certain nombre de vaches malades (présentant par exemple des boiteries), ont été classées comme saines par les éleveurs indépendamment de leur statut vis-à-vis des alertes de santé (erreurs non différentielles). Les conséquences de ces erreurs de classification sur la force de l'association entre exposition (alertes de santé) et maladie, et sur les sensibilités et spécificités associées sont difficilement estimables et corrigibles (Bouyer et al., 2010) La détection par l'éleveur de certains troubles comme les boiteries ayant été montré très peu sensible, **il convient même parfois de se demander si l'outil de monitoring ne doit pas être utilisé comme la méthode de référence pour classifier l'état de santé des animaux**. Les outils de monitoring deviennent alors des outils permettant d'évaluer la sous-détection des éleveurs (Timsit et al., 2011b) concernant une ou plusieurs maladies d'élevage. Dans le cas de Heatime-HR® cependant, on peut supposer que les grandeurs mesurées (activité générale et durées de rumination) sont impactées par de trop nombreux facteurs (alimentation (Adin et al., 2009), management du troupeau (Schirmann et al., 2011), stress thermique (Soriani et al., 2013)...) pour qu'un outil basé sur ces mesures puisse être utilisé comme méthode de référence pour évaluer l'état de santé des vaches laitières. Dans le **chapitre 3.III**, nous avons d'ailleurs constaté que par rapport à la méthode de référence la plus fiable dont nous disposons (examen clinique systématique des vaches), les alertes de santé ne pouvaient pas présenter à la fois une spécificité et une sensibilité élevées.

1.5 Complémentarité de l'activité et de la rumination pour la détection des troubles de santé

Dans le **chapitre 3.I**, nous avons également mis en évidence la complémentarité entre les 2 grandeurs mesurées par Heatime-HR® pour la détection des troubles de santé. Pour une spécificité donnée, la sensibilité du CUSUM sur l'activité ou la rumination était toujours bien supérieure à la sensibilité de chacun des 2 CUSUM pris séparément, indiquant que certains troubles de santé étaient détectés uniquement par le CUSUM sur la rumination et d'autres uniquement par le CUSUM sur l'activité. Cependant, le gain en sensibilité (+10 points avec le réglage conseillé du CUSUM) associé à la combinaison des CUSUM sur l'activité ou la rumination s'accompagnait également d'une perte de spécificité (-5 points). Le gain en termes de sensibilité représente 41 troubles de santé supplémentaires détectés sur l'ensemble de l'étude (171 vaches suivies pendant 2,5 ans) en combinant les 2 CUSUM, par rapport à la situation où seul le CUSUM sur l'activité est utilisé. Tous les types de troubles de santé sont dans ce cas un peu mieux détectés.

Un autre résultat intéressant de ce même chapitre est que pour une spécificité identique, la sensibilité du CUSUM appliqué aux données d'activité était égale ou un peu supérieure à celle du CUSUM appliqué aux données de rumination. Or la mesure la plus innovante du système Heatime-HR®, qui le distingue des systèmes de détection automatisée « classique » des chaleurs et fait qu'il est vendu comme un outil pouvant potentiellement détecter des troubles de santé, est précisément la mesure en continu des durées de rumination. Il était donc attendu que le CUSUM sur la rumination soit plus sensible que le CUSUM sur l'activité, à spécificité comparable. Le résultat inverse obtenu montre que l'activité pourrait donc être une grandeur d'intérêt également pour la détection de maladies autres que les boiteries. Une baisse significative de l'activité (mesurée par des accéléromètres et des podomètres), a d'ailleurs été observée chez des veaux suite à l'inoculation de bactéries *Mannheimia haemolytica*, responsables de bronchopneumonies chez les bovins (Hanzlicek et al., 2010).

2. Généralisation des résultats

La généralisation des résultats concernant l'absence d'intérêt de la mesure de la rumination pour la prédiction de l'ingestion ayant été traitée dans le chapitre 4-1.1, nous traiterons ici que de la généralisation des résultats concernant les alertes de santé.

2.1 Une estimation de l'ordre de grandeur des valeurs informatives des alertes de santé

En raison du coût de la récolte des données (équipement des élevages en outils de monitoring notamment) et du volume de données à traiter, nous avons évalué les performances de détection des algorithmes développés dans 3 fermes uniquement. Or les performances de détection des algorithmes sont variables d'une ferme à l'autre. Nous avons par exemple montré (chapitre 3.II) que 38% des chutes majeures de production laitière donnaient lieu à des alertes de santé dans une ferme contre seulement 20% dans l'autre ferme de l'étude. La spécificité estimée est également probablement variable entre fermes. La variabilité des performances entre fermes a été démontrée dans d'autres travaux, par exemple par Kamphuis et al. (2010) pour un modèle de détection des mammites. Pour une spécificité identique de 99% par exemple, la sensibilité de leur modèle de détection variait de 0 à 71% suivant l'exploitation (9 exploitations). Ces auteurs attribuaient la variabilité observée à la variabilité entre éleveurs dans l'organisation de la détection des mammites, la méthode de référence de l'étude étant la détection par l'éleveur, assisté de listes générées par son robot de traite. Une partie de la variabilité observée peut néanmoins être attribuée également à la variabilité des conditions d'élevage entre exploitations (management du troupeau, alimentation, variabilité de l'expression des maladies, statut sanitaire du troupeau, conditions climatiques...). Cette variabilité peut impacter à la fois la sensibilité des détections (niveau d'expression des boiteries en fonction du bâtiment d'élevage par exemple) que la spécificité (faux positifs liés à des changements d'alimentation provoquant des chutes de rumination). Malgré cette variabilité importante, il est courant que l'estimation de valeurs informatives de systèmes de monitoring de la santé soit limitée à une ou quelques exploitations (Rutten et al., 2013a). Une solution, pour permettre en partie une meilleure adaptation des performances à chaque exploitation, serait de permettre un réglage personnalisé par l'éleveur et ses conseillers, des

seuils de déclenchement des alertes de santé, à la manière de ce qui est observé pour les robots de traite (cf chapitre 4 -4.1 et (Dreno, 2009)).

Idéalement, le développement d’algorithmes se fait sur un jeu de données indépendant du jeu de données utilisé pour leur évaluation, soit en séparant en 2 le jeu de données (Huybrechts et al., 2014), soit en appliquant une procédure de validation croisée (Van Hertem et al., 2013). Dans le premier cas, l’entraînement du modèle (calcul des paramètres du modèle de prédiction par exemple) est effectué sur la première moitié du jeu de données, et l’évaluation du modèle, avec calcul de la sensibilité et de la spécificité par exemple, est effectuée sur la deuxième moitié du jeu de données. Dans le cas de la validation croisée, le jeu de données est généralement découpé aléatoirement en 10 sous-échantillons égaux, avec un ratio égal de périodes cas (vaches malades) et témoins (vaches non malades) dans chaque sous échantillon. Neufs sous-échantillons sont ensuite utilisés pour l’entraînement du modèle, et le dixième sous-échantillon est utilisé pour l’évaluation du modèle. Cette étape est répétée jusqu’à ce que tous les sous-échantillons aient été utilisés une fois pour la validation du modèle. Ce type de méthode permet d’estimer de manière optimale la valeur intrinsèque d’un modèle de détection. Cependant, dans le **chapitre 3.I**, nous avons choisi de ne pas utiliser cette méthode car la procédure de validation était complexe à mettre en place et aurait nécessité un temps de calcul important avec le logiciel d’analyse utilisé (SAS 9.2, SAS Institute Inc., Cary, NC). Au vu de la variabilité importante des valeurs informatives attendues lors de l’application des algorithmes en ferme (cf paragraphe ci-dessus), il nous semble que le gain de précision qui aurait résulté de l’utilisation d’une procédure de validation croisée, sur un jeu de données issues uniquement de 2 fermes, n’aurait pas été majeur.

Trois des études de cette thèse (chapitres 2, 3.I et 3.II) ont été menées en fermes expérimentales. Ce cadre de travail permettait un enregistrement optimal des informations nécessaires aux différentes études. En effet, le personnel des fermes expérimentales est habitué à enregistrer de nombreuses informations au cours des essais qui y sont effectués chaque année, et ces fermes sont souvent équipées de dispositifs de mesure individuels (ingestion, production laitière) qui présentaient un intérêt dans le cadre de cette thèse. La qualité de détection des troubles de santé ainsi que des enregistrements y est donc probablement plus importante que dans la plupart des fermes commerciales. Nous avons par

exemple pu constater que les troubles de santé étaient mieux enregistrés dans les 2 fermes expérimentales ayant participé à cette thèse que dans 5 fermes commerciales dont les données de santé devaient également être analysées initialement, mais qui n'ont pas été utilisées en raison d'enregistrements sanitaires jugés trop aléatoires (5 fois moins de troubles de santé enregistrés par animal présent, pour des situations sanitaires comparables). En ferme commerciale, où la détection des troubles de santé pourrait être parfois moins rigoureuse que dans les fermes expérimentales, l'intérêt d'un dispositif générant des alertes de santé pourrait être supérieur à celui que nous avons trouvé dans cette thèse. La qualité de la détection et de la notation des troubles de santé peut bien sûr influencer sur les valeurs informatives calculées (Kamphuis et al., 2010). Dans cette publication, les fermes dans lesquelles les procédures de détection étaient les plus rigoureuses obtenaient de meilleures performances de détection des mammites.

Les résultats de nos études pourraient être biaisés par des particularités liées aux fermes expérimentales. Cependant, les conditions d'élevage (logement, management de l'alimentation, traite...) dans les 2 fermes expérimentales ayant participé sont relativement proches de celles rencontrées en fermes commerciales. Les résultats que nous avons obtenus concernant la détection des troubles de santé sont donc probablement assez facilement généralisables à ceux qui auraient été obtenus en fermes commerciales. Seul un lot de 60 vaches de la ferme des Trinottières était en essai d'alimentation avec mesure de l'ingestion individuelle quasiment en permanence (sauf juillet et août au pâturage) au cours de la récolte des données, et pouvait donc être géré de manière un peu différente de ce qui se pratique en ferme commerciale (pas d'accès au cornadis de 10h à 13 h, au pâturage uniquement 2 mois en été par exemple).

Pour toutes ces raisons, les estimations des valeurs informatives des alertes de santé développées au cours de cette thèse sont probablement imprécises, mais les tendances observées (sensibilité assez élevée pour la détection des troubles avec atteinte de l'état général par exemple) peuvent être analysées, d'autant qu'elles sont parfois concordantes entre études (faible sensibilité tous troubles confondus obtenus dans les chapitre 3.I, 3.II et 3.III).

2.2 Généralisation à d'autres systèmes de production

La mesure des durées de rumination ayant été démontrée fiable dans d'autres races que les Holstein comme les Jersiaises (Byskov et al., 2014), de format très différent, **il est probable que des résultats comparables aux nôtres seraient obtenus avec d'autres races de vaches laitières**. Cependant des différences dans les systèmes d'élevage (système basé sur l'herbe versus zéro pâturage, niveau de production des vaches, traite automatisée...) pourraient faire varier les résultats obtenus en termes de détection des troubles de santé. La taille de troupeau, le niveau de spécialisation et d'intensification de l'exploitation pouvant avoir, dans certains cas extrêmes, une influence sur l'incidence de certains troubles de santé (Fourichon et al., 2001), il est également attendu que les valeurs informatives des alertes de santé dépendent de ces facteurs. En particulier, une valeur prédictive positive supérieure des alertes est attendue dans les systèmes où l'incidence des troubles est plus importante, ou lors de fusions de troupeaux, qui peuvent constituer des périodes à risque pour l'état sanitaire des animaux.

Pour l'instant, l'utilisation du système de monitoring Heatime-HR® en élevage de vaches allaitantes est quasiment impossible du fait de limitations techniques. Les capacités de stockage d'information limitées des capteurs imposent que les données soient transférées au moins une fois par jour à l'unité centrale Heatime-HR®, ce qui est possible si les animaux viennent au moins une fois par jour au bâtiment (pour la traite). Des solutions techniques existent à la marge pour des animaux au pré qui seraient situés près des bâtiments mais ces solutions ne sont pas adaptées à la grande majorité des élevages allaitants. En admettant que ces limites techniques soient solutionnées à l'avenir, ou que le prix de l'équipement baisse suffisamment pour équiper les vaches allaitantes durant la période hivernale au bâtiment, il est difficile d'extrapoler les résultats obtenus dans cette thèse. En effet, les maladies rencontrées chez les vaches allaitantes sont très différentes de celles rencontrées chez les vaches laitières (dystocies, rétentions placentaires, renversement de matrice versus mammites, boiteries, troubles métaboliques). Pour une même maladie, les expressions cliniques peuvent également être différentes entre les vaches laitières et allaitantes.

Un système comme le Heatime-HR® pourrait techniquement être utilisé pour la détection des troubles de santé chez des jeunes bovins à l'engraissement. Cependant, il a été montré que la mesure des durées de rumination était très imprécise chez des taurillons (Goldhawk et al., 2013), ce qui peut détériorer la performance d'alertes basées sur cette mesure. Les maladies

rencontrées de manière très majoritaire en atelier d'engraissement sont les troubles respiratoires (Smith, 1998), pour la détection desquels nous n'avons pas évalué les alertes de santé développées au cours de cette thèse. D'autre part, des systèmes de mesure automatisée de la température permettent de générer des alertes sensibles (Timsit et al., 2011b) et précoces (Timsit et al., 2011a) pour la détection des troubles de santé, ce qui en fait probablement les systèmes de monitoring à privilégier pour ce type d'élevage.

Globalement, les résultats obtenus sont probablement généralisables à d'autres races de vaches laitières que les Holstein, mais difficilement à d'autres systèmes de production en raison de limitations techniques pour la mesure des données d'activité et de rumination, bien que les technologies de l'élevage de précision se développent également en système extensif désormais (Bocquier et al., 2014, Rutter, 2014).

2.3 Les algorithmes commerciaux n'ont pas été évalués

Pour plusieurs raisons, nous n'avons pas évalué les algorithmes commerciaux utilisés dans l'appareil Heatime-HR®. En effet, au moment du début de la récolte des données utilisées dans cette thèse, l'appareil Heatime-HR® donnait uniquement accès à l'activité et aux durées de rumination sans qu'aucune alerte de santé ne soit générée à partir de ces données. Les fermes expérimentales participant à la thèse avaient donc été équipées avec un système ne fournissant pas les alertes commerciales. De plus, les alertes commerciales sont la propriété du fabricant et nous n'y avons donc pas accès. Il nous aurait donc été impossible de tester différents réglages de sensibilité et spécificité si l'on choisissait de travailler avec les algorithmes commerciaux. De plus, les algorithmes commerciaux générant les alertes de santé étaient susceptibles d'être modifiés régulièrement au cours des 3 années de thèse, de nouveaux développements étant proposés régulièrement par le fabricant du Heatime-HR®. L'évaluation des alertes nécessitant un suivi long des élevages compte tenu de la faible incidence des troubles cliniques, il aurait été difficile d'évaluer les algorithmes commerciaux. Enfin, suivant l'intérêt et les performances des algorithmes développés au cours de cette thèse, il aurait été possible de valoriser ces algorithmes en les proposant au fabricant de l'outil Heatime-HR® au cours d'une négociation commerciale. L'avantage de la méthode choisie est que les résultats sont reproductibles et que les algorithmes utilisés sont connus. Cependant, ce choix de travailler avec des algorithmes développés indépendamment des algorithmes

commerciaux fait que les résultats obtenus ne permettent pas de se prononcer sur les performances des algorithmes commerciaux. Plus généralement, il est possible que l'utilisation d'autres algorithmes ou d'autres traitements statistiques des données puisse conduire à de meilleurs résultats pour la détection des troubles de santé à partir de l'activité et des durées de rumination (cf chapitre 4 -4.4.3).

3. Originalités et limites de la stratégie d'analyse

3.1 Une évaluation d'emblée multi-maladies

Contrairement à certaines études qui ont évalué les performances d'un système de détection vis-à-vis d'une seule maladie comme les boiteries (Van Hertem et al., 2013), la méthodologie que nous avons employée nous a permis d'évaluer d'emblée les performances de notre système d'alerte pour la détection des principales maladies de production rencontrées dans les élevages suivis. Ce choix a été fait car de nombreux troubles impactaient potentiellement les durées de rumination et l'activité générale, sans qu'une étude de la bibliographie ne permette de définir *a priori* quels troubles seraient détectables. De plus, l'outil Heatime-HR® est bien vendu et utilisé en élevage comme un outil de monitoring global de la reproduction (détection des chaleurs) et de la santé. L'ambition de cet outil est donc bien de détecter un large panel de troubles de santé, contrairement aux systèmes mesurant la conductivité électrique du lait par exemple, qui sont davantage « spécialisés » dans la détection des mammites. Une évaluation multi-maladies peut conduire à des erreurs de classement pour certains évènements, notamment lorsque 2 maladies sont détectées par la méthode de référence de manière rapprochée dans le temps (une mammite et une boiterie par exemple). Il n'est alors pas certain que l'alerte de santé ou l'absence d'alerte observée pour le système de détection soit à attribuer aux 2 évènements. D'un autre côté, dans l'approche mono-maladie, les vaches ayant eu d'autres types de troubles de santé sont exclues de l'analyse (Adams et al., 2013), afin de s'assurer de pouvoir attribuer une alerte de santé à un trouble détecté par la méthode de référence. Les performances de détection calculées peuvent donc être sous-estimées par rapport à celles qui seraient obtenues en conditions réelles d'élevage, où les vaches peuvent être atteintes de plusieurs maladies de manière rapprochée dans le temps. Notre méthode

permettait donc d'évaluer les alertes de santé dans des conditions proches de l'utilisation réelle et souhaitée des éleveurs.

Une autre originalité de la méthode employée est que nous avons travaillé avec des grilles cliniques permettant d'évaluer la gravité des troubles de santé détectés (chapitres 3.I et 3.III). Même si la reproductibilité, la répétabilité et la précision de ces grilles est discutable, cela nous a permis de mettre en évidence que les troubles cliniques avec atteinte de l'état général étaient détectés avec des performances moyennes à hautes (sensibilité supérieure à 60%). Cela permettait de déterminer si les alertes pouvaient détecter au moins les troubles les plus graves, performance qui est attendue *a minima* par les éleveurs.

3.2 La spécificité des alertes a été privilégiée dans un contexte de faible prévalence des troubles de santé

En élevage laitier, travailler avec des systèmes d'alerte générant un nombre acceptable de fausses alertes est une préoccupation majeure (De LaRue et al., 2012). En conséquence, nous avons choisi de travailler avec des alertes ayant une spécificité suffisante, a minima 90% dans le chapitre 3.I, pour permettre une utilisation pratique de ces alertes, et non pas de choisir le meilleur compromis entre sensibilité et spécificité des alertes, comme il est classiquement pratiqué pour choisir les réglages optimaux d'un test diagnostique. En effet, la prévalence de troubles de santé en élevage bovin laitier est relativement faible et une spécificité trop faible des alertes va conduire à générer beaucoup de faux positifs alors que les événements vrais positifs seront assez peu nombreux, ce qui diminuera la valeur prédictive positive des alertes. Nous avons ainsi trouvé que la valeur prédictive positive des alertes était 2,5 fois moins importante pour une spécificité de 51% par rapport à une spécificité de 90%. Les nombreux faux positifs générés lorsque la spécificité des alertes est insuffisante risquent de rajouter du travail à l'éleveur (examen des animaux en alerte), de diminuer sa confiance dans les informations fournies par le système d'alerte, et d'augmenter sa charge mentale (Hostiou et al., 2014). Ainsi, un travail mené dans 14 élevages équipés de différentes marques de robots de traite (Dreno, 2009) a montré qu'en moyenne, la spécificité approchée des alertes commerciales de détection des mammites était de 92%. Dans ces conditions, en moyenne 8,3 vaches pour 100 vaches présentes étaient en alerte quotidiennement, seules 15% des vaches en alerte de mammite étaient examinées par les éleveurs, et seulement 50% des vaches

examinées recevaient un traitement mammité. Au final, seules 6% des vaches en alerte recevaient un traitement mammité. Ce travail soulignait donc l'imperfection des alertes commerciales de détection des mammites, qui contraignent l'éleveur à trier parmi de nombreux faux positifs pour détecter les vaches à traiter, ce qui représente du temps de travail et une augmentation de la charge mentale. Les critères de ce tri sont, de plus, relativement empiriques et propres à chaque éleveur, et peuvent aboutir à ne pas traiter des animaux en mammité. Ce travail souligne également qu'il peut exister des différences importantes entre les performances « théoriques » d'un système de détection, pouvant être évalué par exemple ici dans un protocole scientifique utilisant comme méthode de référence la détection des mammites par tirage systématique des premiers jets ou par mesure de la concentration en cellules somatiques du lait, et les performances réelles du système en élevage. Celles-ci sont en effet probablement considérablement moins bonnes, puisque dans l'étude de Dreno (2009), seulement 15% des vaches en alerte sont examinées, le statut sanitaire des autres animaux restant indéterminé pour l'éleveur. On peut par ailleurs remarquer que les valeurs trouvées pour les robots de traite (spécificité approchée = 92%, taux de faux positifs = 8,3/j pour 100 vaches présentes) sont proches des valeurs que nous avons calculé dans le chapitre 3.I pour les alertes de santé (spécificité = 90%, taux de faux positifs = 4,0/j pour 100 vaches présentes).

Concernant les algorithmes développés au cours de cette thèse, des réglages de sensibilité et de spécificité différents (sensibilité plus élevée, spécificité plus faible) pourraient être préférés dans des systèmes d'élevage différents, par exemple chez des jeunes bovins à l'engraissement où la prévalence de troubles de santé est plus élevée et où la très grande majorité des troubles de santé est constituée par les atteintes respiratoires. Dans ce contexte, la spécificité peut être plus faible sans trop dégrader la valeur prédictive positive des alertes (prévalence élevée) et l'examen des animaux en alerte peut prendre moins de temps à l'éleveur car quasiment une seule maladie est recherchée à l'examen clinique (bronchopneumonie infectieuse). De plus, les outils de monitoring peuvent être bien adaptés à ce type d'élevage, où l'observation des animaux peut être moins régulière (traite 2 fois par jour en vaches laitières) et où la période de risque vis-à-vis des bronchopneumonies est limitée dans le temps (après l'allotement).

3.3 Les évènements physiologiques pouvant impacter activité et rumination ont été pris en compte dans l'analyse des données

Contrairement à certaines études où les perturbations liées à certains évènements physiologiques comme les vêlages sont exclues de l'analyse ou incomplètement prises en compte (Adams et al., 2013), nous avons tenu compte au moment de l'analyse de 2 évènements physiologiques majeurs pouvant impacter l'activité et les durées de rumination. Les performances des algorithmes CUSUM ont été évaluées spécifiquement autour du vêlage (chapitre 3.I), permettant de mettre en évidence une baisse de la spécificité des alertes de santé en *peripartum* (baisse de rumination et d'activité observées physiologiquement au moment du vêlage). Les chaleurs sont également souvent associées physiologiquement à des baisses de rumination (Reith et Hoy, 2012). Les alertes de santé concomitantes à des hyperactivités liées aux chaleurs n'ont donc pas été considérées comme des vrais positifs mais ont été ignorées lors de l'analyse, pour éviter d'attribuer des baisses de rumination liées aux chaleurs à des troubles de santé. L'éleveur peut en effet ne pas tenir compte d'une vache en alerte de santé présentant une baisse de rumination si elle est simultanément dans la liste des vaches en alerte de chaleur.

3.4 Le CUSUM : une méthode robuste mais où la prédiction des valeurs attendues d'activité et de rumination s'est révélée difficile

Il n'existe pas de méthode pour choisir *a priori*, en fonction des caractéristiques des données à traiter, la ou les meilleures méthodes, potentiellement les plus performantes pour détecter des déviations de données biologiques continues. Nous avons choisi d'utiliser un test CUSUM car cette méthode a été utilisée de nombreuses fois pour la surveillance de processus biologiques chez les animaux de production (De Vries et Reneau, 2010), et avait donné de bons résultats pour la détection d'hyperthermies chez des jeunes bovins, dans des études menées précédemment dans notre laboratoire (Timsit et al., 2011a, Timsit et al., 2011b). Ce type de test est adapté à la détection de changements assez limités du processus étudié ($\leq 1,5$ écart-type) (Huybrechts et al., 2014). L'utilisation d'un test CUSUM comporte 2 étapes : une étape de prédiction de la valeur attendue pour la variable d'intérêt, et une étape d'application du test CUSUM sur les résidus de la prédiction. Lors des étapes préliminaires de développement des

algorithmes CUSUM, plusieurs modèles ont été testés pour tenter de prédire au mieux l'activité et les durées de rumination des vaches : modèles autorégressifs, modèles linéaires mixtes, utilisation de fonctions sinusoïdales, prise en compte des variations d'activité et de rumination observées dans le reste du troupeau pour la prédiction individuelle...Aucune de ces stratégies de prédiction n'a permis de prédire de manière très fine l'activité ou la rumination des animaux. Finalement, un modèle acceptable et robuste (notamment à la présence de valeurs manquantes fréquentes, en raison des incidents techniques rencontrés sur les capteurs) a été obtenu en prédisant les valeurs attendues à l'aide des valeurs observées à la même heure de la journée lors des jours précédents, une bonne répétabilité de ces valeurs ayant été observée (cf chapitre 4 -1.1).

La prise en compte de la corrélation existant entre les séries temporelles d'activité et de rumination, par exemple en utilisant un CUSUM multivarié (Miekley et al., 2013), aurait pu permettre d'améliorer les performances de détection des algorithmes développés.

4. Perspectives

Les perspectives ouvertes par ce travail sont de plusieurs ordres.

4.1 Perspectives pour les éleveurs

Concernant les alertes de santé, il semble que de meilleures performances de détection que celles des alertes développées au cours de cette thèse doivent être obtenues pour que les éleveurs puissent se reposer davantage sur un tel outil de monitoring global de la santé en élevage bovin laitier.

Ce travail a permis d'obtenir des estimations des sensibilités obtenues par les alertes de santé développées, pour la détection des principaux troubles de santé rencontrés en élevage. Ces résultats montrent que les algorithmes développés détectent avec une sensibilité moyenne à élevée les troubles de santé avec atteinte de l'état général (mammites avec atteinte de l'état général, atteintes de l'état général sans cause identifiée). D'autre part les troubles rares mais

graves (déplacement de caillette à gauche, suspicion de réticulopéritonite traumatique) étaient généralement détectés avec les réglages CUSUM conseillés dans le **chapitre 3.I**. Cela montre que les éleveurs pourraient en règle générale s'appuyer sur les alertes de santé pour la détection des troubles les plus graves. Cela peut être important lorsque l'appareil sert à seconder la personne en charge de la surveillance du troupeau, si celle-ci n'est pas celle qui a habituellement cette charge (garde tournante au sein des Gaec par exemple). Les résultats signifient également que les éleveurs ne pourraient se reposer entièrement sur les alertes de santé développées pour la détection de la majorité des troubles de santé observés en élevage (mammites, boiteries). Ces alertes de santé seraient donc des informations complémentaires, l'utilité principale de l'outil restant la détection des chaleurs, qui elle a été démontrée fiable (cf chapitre 1). Il faut néanmoins signaler que les performances des alertes de santé commerciales sont inconnues, mais que ces alertes font l'objet de mises à jour régulières de la part du fabricant, avec une amélioration possible des algorithmes d'alerte. Les performances de l'outil pourraient donc s'améliorer avec le temps.

La possibilité d'alertes précoces, survenant 1 à 4 jours avant l'apparition des signes cliniques, doit également être expliquée aux éleveurs lors de la prise en main du système d'alerte. La conduite à tenir est de maintenir une surveillance approfondie des vaches en alerte pendant 3 à 4 jours après le début de l'alerte, afin de surveiller l'apparition de signes cliniques. Un affichage sur l'écran Heatime-HR® des alertes des 3 derniers jours pourrait par ailleurs permettre aux éleveurs de mieux surveiller leurs animaux.

De plus, toute une partie des usages et intérêts du système de monitoring n'a pas été évaluée au cours de cette thèse : surveillance de la guérison d'après les valeurs d'activité et de rumination observées après un traitement, surveillance des durées de rumination pour la prévention des maladies en *peripartum*, gestion de la santé et de l'alimentation au niveau du troupeau grâce à la visualisation de l'activité et de la rumination moyenne du troupeau ou de certains lots, évaluation du bien-être animal... Plus généralement, les usages qui sont faits des outils de monitoring peuvent être différents des usages pour lesquels ils avaient été pensés et ces nouveaux usages ou usages détournés peuvent être une source d'amélioration des usages existants. L'association des éleveurs à la conception et au développement de ces outils devrait donc être favorisée.

Tout au long de ce travail de thèse, un soin particulier a été apporté à discuter de l'applicabilité en élevage des algorithmes développés et à discuter de l'intérêt des alertes de

santé générées dans le système d'information de l'éleveur. Ainsi, la fréquence des chutes de production laitière détectables et détectées (**chapitre 3.II**) ont été calculées. De même, des statistiques simples (taux de faux positifs, valeurs prédictives) permettant de rendre compte de l'applicabilité des algorithmes en ferme ont été calculées. Les seuils des algorithmes commerciaux étant paramétrables, il pourrait être souhaitable que les éleveurs, aidés du distributeur de l'outil de monitoring, essayent d'adapter les seuils des algorithmes de santé s'ils ne sont pas satisfaits du paramétrage par défaut de l'outil (remonter les seuils en cas de faux positifs en excès par exemple). Cela pourrait permettre aux éleveurs de s'approprier davantage les informations fournies par l'outil, et d'adapter le réglage des alertes à la situation sanitaire de l'élevage (incidence de troubles de santé) et aux objectifs poursuivis (taux de faux positifs acceptable par exemple). Une telle appropriation peut sembler complexe, mais les éleveurs équipés de robots de traite par exemple paramètrent parfois de manière empirique et personnalisée les seuils d'alertes de mammites (Dreno, 2009).

Les résultats du chapitre 3.I montrent également qu'aucune des principales maladies d'élevage n'est nettement mieux détectée par le CUSUM sur la rumination que par le CUSUM sur l'activité. On pouvait par exemple s'attendre à ce que les boiteries soient nettement mieux détectées par le CUSUM sur l'activité, et les troubles digestifs nettement mieux détectés par le CUSUM sur la rumination. Or ce ne sont pas les résultats que nous avons obtenu. Pour l'éleveur, cela signifie que la nature de l'alerte de santé déclenchée ne permet pas d'orienter le diagnostic vers une ou vers quelques maladies. En présence d'une alerte de santé chez une vache, et en l'absence d'informations complémentaires pouvant orienter le diagnostic (observations visuelles, baisses de production laitière par exemple) l'éleveur devra donc réaliser un examen clinique complet de l'animal pour déterminer son état de santé, ou s'il le juge nécessaire, appeler son vétérinaire. Les compétences de l'éleveur en matière de détection et de traitement des troubles de santé ne pourront donc en aucun cas être moins importantes s'il investit dans un outil de monitoring comme Heatime-HR®, mais elles devront être maintenues voire renforcées.

4.2 Perspectives pour les organismes de développement agricole et les conseillers d'élevage

Les gains de productivité attendus liés au déploiement des technologies de l'élevage de précision sont importants (Bewley, 2013, Wathes et al., 2008). Cependant, ces technologies sont relativement nouvelles et leur mise en place progressive et optimale en élevage nécessite que les fabricants, distributeurs, conseillers en élevage et organismes de développement agricole s'impliquent dans leur déploiement.

A un premier niveau, l'investissement réussi dans une technologie de l'élevage de précision nécessite un conseil personnalisé de l'éleveur au moment de l'équipement. Ces technologies font effet appel à des outils de haute technologie, dont la prise en main et l'appropriation sont des processus progressifs, qui peuvent ne pas être bien vécus par tous les éleveurs (Eastwood, 2008). Il convient donc d'étudier quels sont les moteurs et les freins à l'adoption de ces technologies chez les éleveurs. Quelques travaux ont cherché à déterminer les **raisons pour lesquelles les éleveurs choisissaient de s'équiper en outils de monitoring**. Dans l'étude réalisée par Courties et al. (2014) dans 20 élevages équipés de détecteurs automatisés des chaleurs, 3 motivations principales à l'équipement sont évoquées par les éleveurs : l'amélioration des performances de détection des chaleurs (13/20), le gain de temps pour la détection des chaleurs (5/20) ou l'amélioration d'un autre point de confort de travail (2/20). L'impact de l'équipement en termes de confort de travail et la satisfaction des éleveurs étaient également évalués dans cette étude. L'impact positif des outils de détection des chaleurs sur le stress au moment de la prise de décision de l'insémination est par exemple souligné par 17 éleveurs sur les 20 enquêtés. L'information apportée par l'outil leur permet de prendre une décision plus rapidement et plus sereinement, puisqu'ils peuvent confronter leur opinion à d'autres informations jugées objectives. Une partie de la charge mentale liée à la prise de décision est donc « déléguée » à la machine en quelque sorte. Parmi les 20 éleveurs enquêtés, seuls 2 étaient globalement mécontents de leur équipement, en raison de performances techniques de l'outil jugées insuffisantes. Cela souligne l'importance d'évaluer correctement les outils au moment de leur développement, et de mettre sur le marché des outils qui soient suffisamment performants (cf chapitre 4 -4.3). On peut supposer par ailleurs que, à l'image d'autres outils technologiques, les outils de monitoring mobilisent des techno-imaginaires très forts autour de l'abolition du temps et de l'espace (contrôle de l'exploitation en continu) par exemple, ce qui explique en partie leur succès. A notre connaissance, une seule étude

(Russell et Bewley, 2013) a tenté d'identifier les **principaux freins à l'investissement dans les technologies de l'élevage de précision chez les vaches laitières**. Dans cette étude, réalisée auprès de 236 éleveurs laitiers du Kentucky (USA), les 5 premiers freins évoqués par les éleveurs étaient la méconnaissance des technologies disponibles (55% des éleveurs), un ratio cout/bénéfice jugé trop élevé (42% des éleveurs), trop d'informations fournies par les outils sans savoir quoi en faire (36% des éleveurs), pas de temps à passer sur des outils technologiques (31% des éleveurs), et un manque d'intérêt d'un point de vue économique (30% des éleveurs). Même si les déterminants de l'investissement et de l'adoption des technologies de l'élevage de précision sont nombreux (âge, niveau d'éducation, objectifs, taille d'exploitation, réseau professionnel...), il semble que les éleveurs adopteront plus facilement ces technologies si les bénéfices obtenus suite à l'adoption sont clairs et que la technologie fournie est simple d'utilisation. Pour autant, les moteurs à l'adoption de ces technologies sont puissants (augmentation de la taille des troupeaux, techno-imaginaires, demande de souplesse dans l'organisation des astreintes, diminution des coûts des technologies...) et il est très probable que ces technologies vont entrer encore plus en élevage (Wathes et al., 2008), modifiant profondément le travail des éleveurs (Hostiou et al., 2014). Nous sommes donc encore au début de l'histoire et un champ de recherche peu exploré est disponible pour les sciences humaines et sociales.

A un deuxième niveau, le succès de ces technologies pourrait également passer par un effort dans la valorisation des données issues de capteurs à l'échelle de l'élevage. Cela peut passer par la mise à disposition de l'éleveur, à une fréquence donnée, de tableaux de bord lui permettant de suivre les performances de son troupeau, à la manière de ce qui est fourni par le contrôle laitier par exemple. A notre connaissance, des systèmes d'alarme permettant de détecter de manière automatisée, à partir des mesures individuelles, une dérive de performance du troupeau sont aujourd'hui peu développés sur les outils de monitoring. L'avantage majeur de tels systèmes est que les mesures de monitoring sont quasiment continues, permettant en théorie une détection plus précoce des dérives de performance que les systèmes classiques où la mesure de la production laitière et de la concentration en cellules somatiques est faite à une fréquence mensuelle par exemple. On peut imaginer que des indicateurs simples (taux d'alertes de santé journalier ou hebdomadaire, pourcentage de vaches non cyclées au-delà de 50 jours par exemple) pourraient permettre de suivre efficacement les performances de santé et de reproduction de l'élevage et de détecter

précocément un épisode de grippe intestinale ou une anomalie dans le bilan énergétique des vaches par exemple.

Le troisième niveau de valorisation des données est la valorisation collective des données issues de capteurs pour l'évaluation génétique des animaux. Elle est développée ci-dessous dans le chapitre 4 -4.4.5. Cependant, la valorisation collective des données posera également des problèmes concernant la propriété des données, l'évaluation des outils utilisés, l'échange de données (formats, stockage, accès aux données, échanges entre fabricants) et leur analyse (standardisation des données, méthodologies d'analyse). Un travail pour permettre une identification optimale des animaux et la standardisation des données issues de capteurs est mené au niveau international, notamment par des instances comme le Comité International pour le Contrôle des Performances en Élevage (ICAR). Plusieurs années seront cependant probablement nécessaires pour aboutir à une valorisation collective des données issues de capteurs.

Enfin, les outils de monitoring sont une formidable source de données pour les conseillers en élevage (inséminateurs, vétérinaires, conseillers en nutrition, en santé...). Les personnes qui auront accès à ces données et seront capables de les utiliser pour améliorer la qualité de leur conseil seront en position de force pour conseiller l'éleveur et lui vendre des services. On peut par exemple imaginer des conseillers qui puissent suivre à distance des élevages grâce aux données issues de capteurs, et qui soient capables de prévenir précocément l'éleveur lorsque la santé ou les performances de reproduction de son élevage sont dégradées, en proposant également des mesures d'intervention adaptées (traitements médicamenteux, vaccination...). Les mesures comportementales (activité, rumination) ou physiologiques (température corporelle) réalisées par les capteurs pourront par exemple être utilisées par les vétérinaires et intégrées comme autant d'informations permettant de poser un diagnostic clinique, ouvrant le pas à une médecine de précision. Cependant, la nouvelle finesse de mesure permise par ces capteurs (mesures en continu, sur d'avantage d'animaux...) obligera probablement dans certains cas à redéfinir et à redécrire les phénomènes physiologiques et pathologiques (température corporelle « normale » par exemple).

4.3 Perspectives pour les développeurs/évaluateurs d'outils de monitoring

4.3.1 Des outils de détection des troubles de santé perfectibles

Pour l'entreprise Evolution, ce travail de thèse a permis de mieux connaître les potentialités de l'outil Heatime-HR®. Les résultats de cette thèse ainsi que l'étude de la bibliographie montrent que contrairement aux systèmes de détection des chaleurs, qui sont autant voire plus performants que les éleveurs en moyenne (At-Taras et Spahr, 2001, Rutten et al., 2013a), les systèmes automatisés de détection des troubles de santé sont généralement moins performants que les éleveurs. De nombreux faux positifs sont en général générés par ce type de systèmes (Adams et al., 2013, Miekley et al., 2013), même si nous avons vu dans les **chapitres 3.II et 3.III** qu'une partie de ces alertes correspondaient en réalité à des troubles de santé, pas forcément détectés lorsque la méthode de référence était peu sensible.

Plus généralement, plusieurs leviers peuvent être identifiés pour l'amélioration des performances des systèmes de monitoring. Une première piste de travail est **l'amélioration de capteurs existants**. Ainsi, l'arrivée sur le marché d'accéléromètres de dernière génération (MEMS = micro electro-mechanical systems) a permis récemment la mise au point d'outils fournissant avec une précision acceptable les durées d'ingestion des vaches laitières (Bikker et al., 2014). Des améliorations techniques liées à la conception des capteurs (miniaturisation, besoins en énergie, fiabilité, process qualité) peuvent également considérablement améliorer la fiabilité des outils, qui est un problème récurrent pour les capteurs destinés aux bovins. L'amélioration des outils peut également passer par la **création de nouveaux capteurs**. Des systèmes d'imagerie en 3 dimensions développés pour la notation de l'état corporel des vaches (cf annexe générale) sont par exemple aujourd'hui plus performants que la notation par un observateur entraîné. Les technologies de l'analyse d'image automatisée, développées chez les volailles et chez le porc, sont également prometteuses pour la détection des troubles de santé. De même les mesures effectuées par des capteurs dans le bâtiment d'élevage et l'environnement sont probablement appelés à se développer pour la gestion automatisée de l'ambiance par exemple (cf annexe générale), à la manière de ce qui se pratique déjà dans les productions aviaires et porcines. Les progrès effectués dans le **traitement du signal, le développement d'algorithmes ou de méthodes statistiques** peuvent également aboutir à une amélioration des modèles de détection. **L'intelligence des capteurs et leur souplesse de**

programmation est également une piste d'amélioration. Des capteurs reprogrammables à distance (à partir de leur base de communication) existent par exemple, ce qui permet de les faire évoluer lorsque de nouveaux algorithmes sont développés, sans avoir à changer entièrement le capteur. On peut également imaginer un détecteur de vèlage qui permette de mettre à jour les index facilité de vèlage des vaches équipées en fonction d'informations transmises par le capteur (durée du vèlage par exemple). Il pourrait également être possible d'adapter les algorithmes de détection des vèlages à chaque animal de manière dynamique afin de prévenir de manière systématique l'éleveur pour les animaux ayant un index facilité de vèlage faible mais ne pas le prévenir pour les animaux ayant un index facilité de vèlage élevé, qui devraient pouvoir vèler sans aide. Enfin, comme nous l'avons vu dans le **chapitre 3.I, la combinaison des informations issues de plusieurs capteurs** peut également augmenter les performances de détection, si le gain en termes de sensibilité est supérieur à la dégradation de la spécificité engendrée par l'ajout d'information. Un capteur auriculaire donnant accès à la fois à l'activité (détection des chaleurs, santé, durée d'ingestion) et à la température (santé) des vaches a par exemple obtenu de bonnes performances lors d'une évaluation récente (Bikker et al., 2014). Pour les entreprises du secteur, organiser une veille technologique et scientifique efficace est un enjeu majeur au vu du dynamisme du marché du monitoring et du nombre importants de nouveaux produits mis sur le marché chaque année.

La combinaison de données issues de capteurs embarqués sur les animaux et de données issues du système d'information de l'élevage (production laitière individuelle (cf chapitre 3.II), conductivité du lait, pesées, capteurs équipant les bâtiments (cf annexe générale) est également une piste intéressante de progrès. Cependant, l'échange et la synthèse de données est aujourd'hui difficile car les normes utilisées par chaque fabricant sont différentes et l'interopérabilité entre systèmes est peu développée. Le résultat est qu'aujourd'hui, très peu de systèmes intègrent des informations extérieures à l'outil de monitoring (économiques, historiques, observations éleveurs) pour de la prise de décision automatisée (Rutten et al., 2013a). Or une prise de décision automatisée est envisageable et peut être souhaitable pour décharger l'éleveur d'un certain nombre de tâches, par exemple pour appeler automatiquement l'inséminateur en cas de chaleur détectée à un stade physiologique adéquat.

Pour certains chercheurs, outre les challenges techniques à résoudre, plusieurs autres paramètres doivent être pris en compte au moment du développement des technologies de l'élevage de précision afin d'assurer leur succès (Wathes et al., 2008). **Une étude approfondie du marché** doit être effectuée afin de cibler les besoins des éleveurs et les

produits ne fonctionnant pas (impasses techniques ou marketing) doivent être arrêtés tôt dans le processus de développement. Pour les innovations majeures, **une analyse bioéthique** (Mepham, 1996 in (Wathes et al., 2008)) de l'acceptabilité du produit doit être menée. Une telle analyse repose sur une évaluation par les différentes parties prenantes (éleveur, consommateurs, défenseurs des animaux...) des impacts potentiels de l'innovation sur 3 principes éthiques (bien être, autonomie, justice). Une analyse bioéthique menée au sujet des robots de traite (Millar, 2000, Millar et Mepham, 2001 in (Wathes et al., 2008)) a par exemple conclu à une absence de problème éthique majeur posée par leur utilisation.

4.3.2 Une méthodologie réutilisable pour l'évaluation d'autres outils

La méthode, les algorithmes et les programmes développés au cours de cette thèse pourraient servir à évaluer d'autres outils de monitoring, le développement d'outils de monitoring constituant désormais une partie importante de l'activité de Recherche et Développement d'Evolution. Nos résultats ont également permis de montrer que l'évaluation d'un outil de détection des troubles de santé par rapport à la détection par l'éleveur n'était pas très sensible et que d'autres informations (chutes de production laitière, examen clinique vétérinaire) pouvaient être utilisées pour évaluer de manière optimale la sensibilité et pour éviter de sous-estimer la spécificité des alertes de santé (cf chapitre 4- 1.3.3).

De plus en plus d'outils de monitoring combinent des capteurs enregistrant différentes grandeurs (accélérations, température...) (Bikker et al., 2014) et peuvent donc détecter potentiellement plusieurs types de maladies (boiteries, maladies hyperthermisantes...). Pour des raisons de coût de l'évaluation des outils, il faudra donc évaluer lors d'un protocole unique ou de quelques protocoles, les performances de détection de l'outil à évaluer vis-à-vis de plusieurs maladies. Des protocoles simplifiés (méthodes de référence « simples » et bon marché) par rapport aux protocoles proposés pour l'évaluation des performances vis-à-vis d'une seule maladie devront probablement être proposés pour être applicables.

4.4 Perspectives pour la recherche

4.4.1 Proposer une méthode d'évaluation standardisée des outils de monitoring de la santé

Nous avons vu en introduction de cette thèse que les méthodes employées pour évaluer les outils de monitoring de la santé étaient peu standardisées, ce qui compliquait la comparaison des résultats entre études. En effet, les méthodes de référence utilisées, la taille des fenêtres temporelles pour le calcul de la sensibilité et de la spécificité, les statistiques calculées pour évaluer le système d'alerte (taux de faux positifs, valeurs prédictives non calculées systématiquement) sont variables entre études. Sur la base des méthodes et statistiques simples utilisées au cours de cette thèse, il pourrait être intéressant de proposer un guide méthodologique pour l'évaluation des outils de monitoring de la santé, à la manière de ce qui a été proposé par Kamphuis et al. (2013) pour les mammites. Dans cette étude, les auteurs proposent et discutent des méthodes à employer pour (i) la détection des mammites cliniques (ii) la détection des vaches ayant des concentrations en cellules somatiques élevées (iii) la détermination du statut vis-à-vis des mammites en fin de lactation afin de décider d'un éventuel traitement au tarissement. Pour chacun de ces 3 aspects, des Gold Standards, des tests d'évaluation, des indicateurs de performance et des objectifs de performance sont proposés et discutés. Un travail similaire pourrait être mené pour les autres principaux troubles de santé rencontrés en élevage laitier (boiteries, troubles de la reproduction, troubles métaboliques), sur la base de la méthodologie employée dans cette thèse.

4.4.2 Evaluer la valeur informative des alertes de santé à partir des résultats de plusieurs tests imparfaits

Les alertes de santé ont été évaluées au cours de cette thèse par rapport à des méthodes de référence imparfaites c'est-à-dire en l'absence de Gold Standard. Dans ce cas, on peut combiner les résultats obtenus par différentes méthodes de référence imparfaites pour calculer la sensibilité et la spécificité du test diagnostique à évaluer (Dohoo et al., 2010). Une première méthode dite **méthode de référence composite** ('composite reference standard') consiste à appliquer un premier test sur tous les échantillons à tester, puis d'appliquer un second test sur

les échantillons négatifs à l'issue du premier test. Un échantillon est considéré positif s'il est positif par l'un des 2 tests utilisés, et est considéré négatif s'il négatif par les 2 tests. Une autre méthode en absence de Gold standard consiste à utiliser des **modèle à structure latente** ('Latent class model'). Dans ce cas, au moins 2 tests diagnostiques sont appliqués en parallèle et le vrai état de santé des animaux est considéré comme une variable latente inconnue. Sous réserve que (i) la prévalence attendue dans au moins 2 populations de l'étude soit différente (ii) la sensibilité et la spécificité des tests diagnostiques soient constantes dans les différentes populations (iii) les résultats des tests utilisés soient indépendants, alors la sensibilité vraie d'un ou de plusieurs tests diagnostiques peut être calculée (Pouillot et al., 2002). Ces méthodes sont utilisées dans le cas de tests diagnostiques biologiques (sérologies ou PCR par exemple) mais pourraient peut-être être adaptés dans les protocoles d'évaluation d'outils de monitoring. Par exemple dans le chapitre 3.II, nous avons accès aux résultats de 3 tests (détection par l'éleveur, repérage des chutes de production laitière, alertes de santé basées sur l'activité et la rumination) appliqués en parallèle sur les données, dans 2 populations de prévalence différente (vaches malades et périodes témoin). Ces méthodes pourraient permettre d'évaluer plus finement les valeurs informatives vraies des alertes générées par des systèmes de monitoring, pour lesquels les méthodes de référence utilisées sont très généralement imparfaites.

4.4.3 L'utilisation d'autres méthodes mathématiques et statistiques pourraient permettre de mieux détecter les troubles de santé

L'amplitude des baisses de rumination et d'activité attendue lors de trouble de santé étant variable entre les différents types de troubles, et entre les différentes occurrences pour un type de trouble donné, une solution pour détecter davantage de troubles peut être **d'appliquer plusieurs types d'algorithmes simultanément sur les données**. Huybrechts et al. (2014) ont par exemple appliqué sur des données de production laitière journalière à la fois un test cusum, idéal pour détecter les baisses de faible amplitude ($< 1,5$ écart-types), et un test de type Shewhart, idéal pour la détection de baisses de forte amplitude (> 3 écart-types), pour détecter des mammites.

La sensibilité des alertes de santé pourrait également probablement être améliorée en utilisant d'autres méthodes de traitement du signal et/ou d'analyse statistique. Ainsi, une sensibilité

(89%) et une spécificité (85%) élevée ont été obtenues par (Van Hertem et al., 2013) pour un modèle de détection des boiteries, en créant une grande quantité de variables à partir des données d'activité et de rumination issues du capteur HR-Tag (ratio d'activité nuit/jour, activité durant la journée, durée de rumination durant la nuit...). Ces variables étaient ensuite intégrées dans un modèle de détection qui comprenait également des variables liées à la production laitière journalière, très liées à la détection des boiteries. De nombreuses autres méthodes pourraient être utilisées sur les données issues de capteurs : réseaux de neurones (par exemple Sun et al. (2010) pour la conductivité du lait et les infections intramammaires) ; filtres de Kalman (par exemple, pour différents troubles de santé et pour la détection de l'oestrus, de Mol et al. (1999)) ; logique floue (par exemple Kamphuis et al. (2008), également pour les infections intramammaires) ; méthodes de « data mining »...

4.4.4 Des analyses complémentaires pourraient permettre de mieux évaluer les performances de Heatime-HR®

Certains prolongements de ce travail n'ont pas pu être réalisés faute de temps. **La construction d'un modèle statistique afin d'étudier l'effet des différents troubles de santé sur les durées de rumination et sur l'activité aurait pu permettre d'adapter la taille de la fenêtre de validation utilisée pour le calcul de la sensibilité à chaque maladie** en étudiant pendant combien de temps ces grandeurs sont impactées en moyenne avant et après chaque type de trouble de santé, en comparaison avec des données recueillies chez des vaches saines. Lukas et al. (2009) ont par exemple construit un tel modèle pour décrire pendant combien de temps la production laitière et la conductivité électrique du lait étaient impactées avant et après différentes catégories de troubles de santé. Leurs résultats montrent que la production laitière est diminuée significativement 9 jours avant une pneumonie, 6 jours avant une boiterie et entre 1 et 4 jours avant une mammite, suivant sa sévérité, la méthode de référence étant l'observation par l'éleveur. Les auteurs utilisent ensuite une fenêtre unique de 10 jours (10 jours précédents la détection) pour calculer la sensibilité des détections de leur modèle. La taille de cette fenêtre aurait pu être adaptée à chaque trouble de santé (i.e. fenêtre de minimum 6 jours pour les boiteries, et de minimum 4 jours pour les mammites). Une approche similaire aurait pu être utilisée avec les données d'activité et de rumination issues de Heatime-HR®. Cette approche pourrait être particulièrement intéressante pour les troubles

pour lesquels la date de détection par l'éleveur ne correspond très probablement pas à l'apparition des signes cliniques, par exemple pour les boiteries. Ainsi, Van Hertem et al. (2013) ont montré que des variables liées à l'activité et aux durées de rumination de vaches laitières étaient impactée dès 6 à 7 jours avant la détection d'une boiterie par les éleveurs, ce qui encourage à utiliser une fenêtre temporelle de minimum 7 jours pour le calcul de la sensibilité du modèle de détection. Une telle analyse statistique pourrait donner des éléments de réponse pour la taille des fenêtres temporelles à utiliser, même si l'applicabilité des fenêtres doit également être questionnée. Est-il par exemple pertinent d'alerter l'éleveur 7 jours avant une mammite, alors qu'aucun signe clinique n'est détectable ? Quelle action devra-t-il alors accomplir en réponse à cette alerte ? D'autre part est-il logique, compte tenu des connaissances de la physiopathologie des mammites, qu'une alerte de santé déclenchée 7 ou 10 jours avant la détection de grumeaux par l'éleveur soit attribuée à cette mammite ?

Le modèle statistique évoqué précédemment aurait également pu permettre **d'affiner les réglages des CUSUM en fournissant les effets moyens des troubles de santé sur les durées de rumination et sur l'activité**, puisque le CUSUM est optimal pour la détection d'une certaine variation de la grandeur sur laquelle il est appliqué (Luo et al., 2009). Ainsi Huybrechts et al. (2014) ont utilisé l'effet moyen de baisse de production laitière lors de mammite calculé par Lukas et al. (2009), pour trouver les paramètres h et k optimaux de leur détecteur CUSUM.

Un modèle statistique permettant d'analyser les facteurs influençant la sensibilité des détections (saison, parité, proximité de chaleurs...) aurait également permis de compléter les résultats obtenus.

4.4.5 Utiliser les données issues de capteurs pour le phénotypage de caractères de santé

La notion de **phénotypage à haut débit (Hocquette et David, 2011), c'est-à-dire en utilisant les mesures faites en continu par des capteurs sur les animaux pour la mesure de phénotypes d'intérêt**, a été théorisée récemment, par exemple concernant la mesure de la robustesse chez les animaux (Phocas et al., 2014). A notre connaissance très peu d'applications pratiques de cette notion ont abouti pour le moment. Pour les phénotypes déjà mesurés mais qui sont mesurables de manière plus fréquente à moindre frais (note d'état

corporel (cf annexe générale), poids par exemple), les applications en termes de phénotypage pourraient être assez rapides (cf annexe générale). Pour les phénotypes nécessitant une analyse complexe des données pour leur interprétation (spectres moyen infrarouge du lait et composition en acides gras, intensité des chaleurs par exemple), certaines applications pour le phénotypage sont quasiment abouties alors que d'autres pourraient prendre plusieurs années (cf annexe générale). Les performances des détecteurs automatisés des chaleurs étant relativement bonnes, certains projets de recherche visent à utiliser les données issues de ces détecteurs pour **indexer les vaches laitières vis-à-vis de la détectabilité de l'œstrus, et de sa cyclicité**. En effet, face à la forte sélection sur la production laitière des dernières décennies, des défauts d'expression des chaleurs sont apparus de plus en plus fréquemment dans les troupeaux laitiers : une diminution de la durée d'expression des comportements spécifiques des chaleurs (Disenhaus et al., 2010), une expression plus discrète et une cyclicité irrégulière (phase lutéale courte ou au contraire très longue, interruption de cyclicité) impliquant le recours à des traitements hormonaux. Un projet financé par le compte d'affectation spéciale pour le développement agricole et rural, piloté par l'Institut de l'Élevage en collaboration avec l'UMR Bioepar, et intitulé « Monitoring Automatisé de la Reproduction : Innovations et Applications pour l'élevage bovin laitier », doit commencer début 2015. Une partie de ce projet vise à rechercher de nouveaux phénotypes pour caractériser l'expression des chaleurs et la cyclicité, en se basant sur les données issues d'un système de détection automatisée des chaleurs (différent du Heatime-HR®). L'objectif est d'étudier la variabilité et la répétabilité d'indicateurs potentiellement liés à l'intensité des chaleurs (amplitude, durée, aire sous la courbe de la suractivité) ou à la cyclicité des chaleurs (intervalle entre chaleurs, régularité des chaleurs). Cela permettra d'envisager la faisabilité technique et l'opportunité scientifique de développer une évaluation génétique sur l'expression des chaleurs et/ou la cyclicité des vaches en complément de l'index fertilité déjà existant en France. Les résultats de ce projet de recherche, qui utilisera une marque de détecteurs de chaleurs, pourraient être transposables à d'autres outils utilisant des activimètres, dont Heatime-HR®.

4.4.6 Utiliser les données issues de capteurs pour la surveillance syndromique

La surveillance syndromique (Henning, 2004) vise à détecter de manière précoce un large spectre de maladies, à partir de signes non spécifiques d'altération de la santé comme des baisses de production laitière (Madouasse et al., 2014) ou des valeurs anormales d'indicateurs liés aux performances de reproduction (Marceau et al., 2014) par exemple. Des alertes concernant le signalement et la localisation de performances anormales peuvent permettre des investigations et une identification accélérée de problèmes et de leur cause. En santé animale, les méthodes de surveillance syndromique ont été utilisées pour la **détection de maladies émergentes comme la fièvre catarrhale ovine** (Madouasse et al., 2014, Marceau et al., 2014). Il faut d'ailleurs remarquer que les méthodes utilisées pour la surveillance syndromique sont proches de celles que nous avons utilisé au cours de cette thèse (analyse de séries temporelles, tests CUSUM) mais que ces méthodes sont employées au niveau de données issues d'une région et non à l'échelle de la vache. **Les données issues des systèmes de monitoring comme Heatime-HR® pourraient être utilisées pour de la surveillance syndromique.** Des indicateurs liés à la santé (nombre d'alertes de santé par zone géographique par exemple) pourraient être modélisés et les déviations de performance détectées précocement. L'avantage d'utiliser ces données est que la collecte des données est faite en continu sur les animaux. Une limitation pourrait être, pour certains systèmes de monitoring, le faible pourcentage d'élevages équipés pour le moment, ce qui peut aboutir à une couverture incomplète de certaines régions géographiques. Une utilisation à grande échelle nécessiterait également que les données issues de différents systèmes de monitoring de la production laitière par exemple, soient compatibles en termes de formats notamment.

4.4.7 Etudier les conditions techniques et économiques d'un déploiement réussi des technologies de l'élevage de précision

Les résultats de cette thèse pourraient servir à **évaluer l'impact technico-économique d'un dispositif comme le Heatime-HR® en élevage.** La connaissance des performances techniques des outils est en effet une des connaissances nécessaires à l'estimation de l'impact technico-économique d'un outil comme le Heatime-HR®. A notre connaissance, peu de travaux se sont intéressés à l'évaluation de l'impact technico-économique des outils de

monitoring en élevage. Un travail récent de Courties et al. (2014) a évalué l'impact de dispositifs de détection automatisée des chaleurs en élevage laitier, à l'aide d'un simulateur de troupeau. Ce travail concluait qu'investir dans un dispositif automatisé d'aide à la détection des chaleurs était rentable économiquement uniquement dans un grand troupeau (120 vaches) de vaches hautes productrices, ou dans un troupeau de taille moyenne (60 vaches hautes productrices) en vèlages étalés toute l'année (permettant un équipement seulement partiel des animaux en capteurs). L'impact sur le temps de travail n'était toutefois pas comptabilisé dans cette étude. De plus, l'équipement était rentable économiquement pour ce type d'élevage uniquement si la sensibilité de détection des chaleurs était faible avant équipement (40 à 50%), et si l'équipement permettait d'améliorer de 30 à 40 points la sensibilité des détections. D'autres simulations menées dans des troupeaux de 130 vaches (Rutten et al., 2013b) ont également montré l'intérêt économique d'un équipement de détection automatisée des chaleurs. Quelques travaux menés aux Etats-Unis ont abouti à des outils de conseil pour permettre aux éleveurs de faire un choix éclairé lors de l'investissement dans une technologie liée à l'élevage de précision. Ainsi, un simulateur développé par Dolecheck et Bewley (2014) permet de comparer le retour sur investissement lié à l'équipement en détecteur de chaleur (comparaison de 3 outils possible), après description de la ferme, de la gestion de la reproduction, des performances techniques des outils et de leur coût. Un modèle stochastique très intéressant pouvant potentiellement être utilisé pour évaluer l'impact technico-économique des technologies de l'élevage de précision en élevage a également été décrit en détail (Bewley et al., 2010). L'outil développé permet de modéliser l'impact d'un outil de monitoring sur différents paramètres d'élevage (production laitière, réforme, performances de reproduction, incidence des maladies, coûts liés à la santé, performance économique...) Cependant aucun exemple concret d'application à un outil de monitoring n'est décrit dans cet article, peut être justement en partie en raison du manque d'évaluation technique d'outils de monitoring global de la santé comme le Heatime-HR®. D'autre part, il convient de remarquer le taux d'équipement des animaux en capteurs est différent dans le cas des systèmes de détection des chaleurs (taux d'équipement généralement inférieur à 40% compte tenu du nombre de vaches susceptibles d'être en chaleur à moment donné) et des systèmes de détection des troubles de santé, pour lesquels idéalement toutes les vaches doivent être équipées, puisqu'elles sont susceptibles de présenter un trouble de santé. Les simulations d'impact économique doivent tenir compte de cette particularité.

Compte tenu de la complexité des conséquences potentielles de l'usage d'outils de détection automatisée des troubles de santé (conséquences sur le temps de travail, le taux de traitement des animaux, les taux de guérison associés à une détection précoce...), des simulations informatiques en troupeaux virtuels pourraient permettre d'évaluer l'impact technico-économique d'un outil comme Heatime-HR® à la fois sur les performances de reproduction et de santé. Un travail de modélisation doit être mené au sein de l'UMR Bioepar prochainement. Concernant l'impact technico-économique d'un outil de détection des troubles de santé, les valeurs informatives trouvées au cours de ce travail de thèse pourront être utilisées. Cependant, le coût du dispositif étant pour le moment assez élevé pour de bonnes performances de détection des chaleurs, mais des performances de détection des troubles de santé moyennes, il est probable que la rentabilité économique du dispositif ne soit pas atteinte. Néanmoins, cet outil, comme les autres outils de monitoring, peut présenter un intérêt pour les éleveurs par son impact potentiellement positif sur l'organisation et le confort de travail, alors même qu'il n'est pas rentable économiquement.

Références bibliographiques

- Adams, A. E., F. J. Olea-Popelka, and I. N. Roman-Muniz. 2013. Using temperature-sensing reticular boluses to aid in the detection of production diseases in dairy cows. *J Dairy Sci* 96:1549-1555.
- Adin, G., R. Solomon, M. Nikbachat, A. Zenou, E. Yosef, A. Brosh, A. Shabtay, S. J. Mabweesh, I. Halachmi, and J. Miron. 2009. Effect of feeding cows in early lactation with diets differing in roughage-neutral detergent fiber content on intake behavior, rumination, and milk production. *J Dairy Sci* 92:3364-3373.
- At-Taras, E. E. and S. L. Spahr. 2001. Detection and characterization of estrus in dairy cattle with an electronic heatmount detector and an electronic activity tag. *J Dairy Sci* 84:792-798.
- Bareille, N., F. Beaudeau, S. Billon, A. Robert, and P. Faverdin. 2003. Effects of health disorders on feed intake and milk production in dairy cows. *Liv Prod Sci* 83:53-62.
- Bewley, J. M., M. D. Boehlje, A. W. Gray, H. Hogeveen, S. J. Kenyon, S. D. Eicher, and M. M. Schutz. 2010. Stochastic simulation using @risk for dairy business investment decisions. *Agricultural Finance Review* 70:97-125.
- Bewley, J. M. 2013. Exciting dairy breakthroughs : Science fiction or precision dairy farming ? Pages 1-6 in Precision dairy conference and expo, Rochester, Minnesota.
- Bikker, J. P., H. van Laar, P. Rump, J. Doorenbos, K. van Meurs, G. M. Griffioen, and J. Dijkstra. 2014. Technical note: Evaluation of an ear-attached movement sensor to record cow feeding behavior and activity. *J Dairy Sci* 97:2974-2979.
- Bocquier, F., N. Debus, A. Lurette, C. Maton, G. Viudes, C. H. Moulin, and M. Jouven. 2014. Elevage de précision en systèmes peu intensifiés. *INRA Prod Anim* 27:101-111.
- Bouyer, J., D. Hémon, S. Cordier, F. Derriennic, I. Stucker, B. Stengel, and J. Clavel. 2010. *Epidémiologie-principes et méthodes quantitatives*. Ed. Inserm, Paris, France.
- Byskov, M. V., A. K. Schulze, M. R. Weisbjerg, B. Markussen, and P. Norgaard. 2014. Recording rumination time by a rumination monitoring system in jersey heifers fed grass/clover silage and hay at three feeding levels. *J Anim Sci* 92:1110-1118.

Courties, R. 2014. Evaluation multicritères du rapport coûts/bénéfices de l'équipement en capteurs pour la détection automatisée des chaleurs. in Mémoire de fin d'études. Oniris, Nantes.

De LaRue, B., C. Kamphuis, and J. Jago. Accessed Sept. 05, 2014. Smart ways to detect mastitis: What do we know about automated in-line detection systems? http://www.sciquest.org.nz/elibrary/download/82408/Smart_ways_to_detect_mastitis%3A_What_do_we_know_abo.pdf

De Mol, R. M., A. Keen, G. H. Kroeze, and J. M. F. H. Achten. 1999. Description of a detection model for oestrus and diseases in dairy cattle based on time series analysis combined with a kalman filter. *Comput Electron Agr* 22:171-185.

De Vries, A. and J. K. Reneau. 2010. Application of statistical process control charts to monitor changes in animal production systems. *J Anim Sci* 88:E11-24.

Disenhaus, C., E. Cutullic, S. Freret, P. Paccard, and C. Ponsart. 2010. Vers une cohérence des pratiques de détection des chaleurs : Intégrer la vache, l'éleveur et le système d'élevage in *Rencontres Recherche Ruminants*. Paris, France.

Dohoo, I., W. Martin, and H. Stryhn. 2010. *Veterinary epidemiologic research*. 2nd ed., Ed. VER Inc., Charlottetown, Canada.

Dolecheck, K. and J. M. Bewley. Accessed Sept. 05, 2014. Investment analysis of heat detection technologies. <http://afsdairy.ca.uky.edu/HeatDetectionTechnologies>.

Dreno, J. B. 2009. Utilisation des données enregistrées par le robot de traite pour le pilotage de la santé mammaire des vaches laitières. Thèse vétérinaire. Ecole Nationale Vétérinaire de Nantes, Nantes, France..

Eastwood, T. R. 2008. Innovative precision dairy systems: A case study of farmer learning and technology codevelopment. PhD Tesis. The University of Melbourne, Melbourne, Australia.

Fabian, J., R. A. Laven, and H. R. Whay. 2014. The prevalence of lameness on new zealand dairy farms: A comparison of farmer estimate and locomotion scoring. *Vet J* 201:31-38.

- Ferran, A. A., P. L. Toutain, and A. Bousquet-Melou. 2011. Impact of early versus later fluoroquinolone treatment on the clinical, microbiological and resistance outcomes in a mouse-lung model of *pasteurella multocida* infection. *Vet Microbiol* 148:292-297.
- Fourichon, C., H. Seegers, N. Bareille, and F. Beaudeau. 1999. Effects of disease on milk production in the dairy cow: A review. *Prev Vet Med* 41:1-35.
- Fourichon, C., F. Beaudeau, N. Bareille, and H. Seegers. 2001. Incidence of health disorders in dairy farming systems in western france. *Liv Prod Sci* 68:157-170.
- Gernand, E., P. Rehbein, U. U. von Borstel, and S. König. 2012. Incidences of and genetic parameters for mastitis, claw disorders, and common health traits recorded in dairy cattle contract herds. *J Dairy Sci* 95:2144-2156.
- Goldhawk, C., K. Schwartzkopf-Genswein, and K. A. Beauchemin. 2013. Technical note: Validation of rumination collars for beef cattle. *J Anim Sci* 91:2858-2862.
- Hanzlicek, G. A., B. J. White, D. Mosier, D. G. Renter, and D. E. Anderson. 2010. Serial evaluation of physiologic, pathological, and behavioral changes related to disease progression of experimentally induced *mannheimia haemolytica* pneumonia in postweaned calves. *Am J Vet Res* 71:359-369.
- Henning, K. J. 2004. What is syndromic surveillance? *MMWR Morb. Mortal. Wkly. Rep.* 53 (September (Supl.)): 5–11.
- Hillerton, J. E. and J. E. Semmens. 1999. Comparison of treatment of mastitis by oxytocin or antibiotics following detection according to changes in milk electrical conductivity prior to visible signs. *J Dairy Sci* 82:93-98.
- Hocquette, J.-F. and V. David. 2011. Les enjeux du phénotypage. *BTIA Bulletin Technique de l'Insemination Artificielle*: 142.
- Holman, A., J. Thompson, J. E. Routly, J. Cameron, D. N. Jones, D. Grove-White, R. F. Smith, and H. Dobson. 2011. Comparison of oestrus detection methods in dairy cattle. *Vet Rec* 169:47.
- Hostiou, N., C. Allain, S. Chauvat, A. Turlot, C. Pineau, and J. Fagon. 2014. L'élevage de précision : Quelles conséquences pour le travail des éleveurs. *INRA Prod. Anim.* 27:111-120.

Huybrechts, T., K. Mertens, J. De Baerdemaeker, B. De Ketelaere, and W. Saeys. 2014. Early warnings from automatic milk yield monitoring with online synergistic control. *J Dairy Sci* 97:3371-3381.

ISO (International Organization for Standardization). 2007. Automatic milking installations—Requirements and testing. Annex C: Example of methods of evaluating detection systems for milk deemed as abnormal due to blood or to changes in homogeneity. ISO 20966:2007. Geneva, Switzerland.

Kamphuis, C., R. Sherlock, J. Jago, G. Mein, and H. Hogeveen. 2008. Automatic detection of clinical mastitis is improved by in-line monitoring of somatic cell count. *J Dairy Sci* 91:4560-4570.

Kamphuis, C., H. Mollenhorst, J. A. P. Heesterbeek, and H. Hogeveen. 2010. Detection of clinical mastitis with sensor data from automatic milking systems is improved by using decision-tree induction. *J Dairy Sci* 93:3616-3627.

Kamphuis, C., B. Dela Rue, G. Mein, and J. Jago. 2013. Development of protocols to evaluate in-line mastitis-detection systems. *J Dairy Sci* 96:4047-4058.

Lukas, J. M., J. K. Reneau, R. Wallace, D. Hawkins, and C. Munoz-Zanzi. 2009. A novel method of analyzing daily milk production and electrical conductivity to predict disease onset. *J Dairy Sci* 92:5964-5976.

Luo, Y., Z. Li, and Z. Wang. 2009. Adaptive CUSUM control chart with variable sampling intervals. *Comput. Stat. Data Anal.* 53:2693-2701.

Madouasse, A., A. Marceau, A. Lehebel, H. Brouwer-Middelesch, G. van Schaik, Y. Van der Stede, and C. Fourichon. 2014. Use of monthly collected milk yields for the detection of the emergence of the 2007 French btv epizootic. *Prev Vet Med* 113:484-491.

Marceau, A., A. Madouasse, A. Lehebel, G. van Schaik, A. Veldhuis, Y. Van der Stede, and C. Fourichon. 2014. Can routinely recorded reproductive events be used as indicators of disease emergence in dairy cattle? An evaluation of 5 indicators during the emergence of bluetongue virus in France in 2007 and 2008. *J Dairy Sci*. <http://dx.doi.org/10.3168/jds.2013-7346>.

Miekley, B., E. Stamer, I. Traulsen, and J. Krieter. 2013. Implementation of multivariate cumulative sum control charts in mastitis and lameness monitoring. *J Dairy Sci* 96:5723-5733.

Milner, P., K. L. Page, and J. E. Hillerton. 1997. The effects of early antibiotic treatment following diagnosis of mastitis detected by a change in the electrical conductivity of milk. *J Dairy Sci* 80:859-863.

NRC (National Research Council). 2001. Nutrient requirements of dairy cattle. 7th revised edition ed. National academy press. Washington, DC, USA.

Phocas, F., J. Bobe, L. Bodin, B. Charley, J. Y. Dourmad, N. C. Friggens, J. F. Hocquette, P. Y. Le Bail, E. Le Bihan-Duval, P. Mormède, P. Quéré, and F. Schelcher. 2014. Des animaux plus robustes : Un enjeu majeur pour le développement durable des productions animales nécessitant l'essor du phénotypage fin et à haut débit. *INRA Prod. Anim.* 27:181-194.

Pouillot, R., G. Gerbier, and I. A. Gardner. 2002. Tags, a program for the evaluation of test accuracy in the absence of a gold standard. *Prev Vet Med* 53:67-81.

Radostits, O. M., C. C. Gay, K. W. Hinchcliff, and P. D. Constable. 2007. *Veterinary Medicine: A textbook of the diseases of cattle, horses, sheep, pigs, and goats*. 10th ed. Saunders Elsevier, New York.

Reith, S. and S. Hoy. 2012. Relationship between daily rumination time and estrus of dairy cows. *J Dairy Sci* 95:6416-6420.

Russell, R. A. and J. M. Bewley. 2013. Characterization of Kentucky dairy producer decision-making behavior. *J Dairy Sci* 96:4751-4758.

Rutten, C. J., A. G. Velthuis, W. Steeneveld, and H. Hogeveen. 2013a. Invited review: Sensors to support health management on dairy farms. *J Dairy Sci* 96:1928-1952.

Rutten, N., W. Steeneveld, C. Inchaisri, and H. Hogeveen. 2013b. Analysis of investment in an estrus detection system for dairy farms. Pages 177-178 in *Precision dairy conference and expo*, Rochester, Minnesota.

Rutter, S. M. 2014. Smart technologies for detecting animal welfare status and delivering health remedies for rangeland systems. *Rev. Sci. Tech. Off. Int. Epiz.* 33:181-187.

Sannmann, I., O. Burfeind, R. Voigtsberger, and W. Heuwieser. 2013. Comparison of two monitoring and treatment strategies for cows with acute puerperal metritis. *Theriogenology* 79:961-969.

Sannmann, I., O. Burfeind, R. Voigtsberger, and W. Heuwieser. 2014. Erratum to comparison of two monitoring and treatment strategies for cows with acute puerperal metritis • [theriogenology 79 (2013) 961-969]. *Theriogenology* 81:875-876.

Schirmann, K., N. Chapinal, D. M. Weary, W. Heuwieser, and M. A. von Keyserlingk. 2011. Short-term effects of regrouping on behavior of prepartum dairy cows. *J Dairy Sci* 94:2312-2319.

Schirmann, K., N. Chapinal, D. M. Weary, W. Heuwieser, and M. A. von Keyserlingk. 2012. Rumination and its relationship to feeding and lying behavior in holstein dairy cows. *J Dairy Sci* 95:3212-3217.

Smith, R. A. 1998. Impact of disease on feedlot performance: A review. *J Anim Sci* 76:272-274.

Soriani, N., E. Trevisi, and L. Calamari. 2012. Relationships between rumination time, metabolic conditions, and health status in dairy cows during the transition period. *J Anim Sci* 90:4544-4554.

Soriani, N., G. Panella, and L. Calamari. 2013. Rumination time during the summer season and its relationships with metabolic conditions and milk production. *J Dairy Sci* 96:5082-5094.

Sun, Z., S. Samarasinghe, and J. Jago. 2010. Detection of mastitis and its stage of progression by automatic milking systems using artificial neural networks. *J Dairy Res* 77:168-175.

Timsit, E., S. Assie, R. Quiniou, H. Seegers, and N. Bareille. 2011a. Early detection of bovine respiratory disease in young bulls using reticulo-rumen temperature boluses. *Vet J* 190:136-142.

Timsit, E., N. Bareille, H. Seegers, A. Lehebel, and S. Assie. 2011b. Visually undetected fever episodes in newly received beef bulls at a fattening operation: Occurrence, duration, and impact on performance. *J Anim Sci* 89:4272-4280.

van den Borne, B. H., G. van Schaik, T. J. Lam, and M. Nielen. 2010. Therapeutic effects of antimicrobial treatment during lactation of recently acquired bovine subclinical mastitis: Two linked randomized field trials. *J Dairy Sci* 93:218-233.

Van Hertem, T., E. Maltz, A. Antler, C. E. Romanini, S. Viazzi, C. Bahr, A. Schlageter-Tello, C. Lokhorst, D. Berckmans, and I. Halachmi. 2013. Lameness detection based on multivariate continuous sensing of milk yield, rumination, and neck activity. *J Dairy Sci* 96:4286-4298.

Wathes, C. M., H. H. Kristensen, J. M. Aerts, and D. Berckmans. 2008. Is precision livestock farming an engineer's daydream or nightmare, an animal's friend or foe, and a farmer's panacea or pitfall? *Comput Electron Agr* 64:2-10.

Conclusion générale

L'objectif de cette thèse était de documenter l'intérêt de la mesure de l'activité générale et des durées de rumination des vaches laitières par l'appareil de monitoring Heatime-HR®. Les résultats attendus étaient d'explorer les potentialités de l'appareil pour le phénotypage de caractères d'intérêt (prédiction de l'ingestion notamment) et d'évaluer la valeur informative d'alertes générées à partir de l'activité générale et des durées de rumination pour la détection des troubles de santé chez les vaches laitières.

Les principales conclusions de cette thèse sont que :

- La connaissance des durées de rumination ne permet pas de mieux prédire les quantités ingérées. L'addition des durées de rumination dans un modèle de prédiction de l'ingestion (NRC 2001) augmentait significativement la qualité de prédiction du modèle. Cependant l'effet des durées de rumination sur l'ingestion (0,031 kg/h) était trop faible pour que leur prise en compte présente réellement un intérêt. Cependant un outil comme le Heatime-HR® pourrait présenter un intérêt pour phénotyper des caractères afin de permettre l'amélioration des performances de reproduction (cyclicité des chaleurs, détectabilité de l'oestrus) ou de santé (robustesse des animaux).
- Les alertes de santé ne se substituent pas à l'observation des troubles de santé par l'éleveur. Les performances des algorithmes 'cumulative sum' développés spécifiquement dans cette thèse pour la détection des troubles de santé à partir des mesures continues d'activité et de rumination sont insuffisantes pour que les éleveurs se reposent entièrement sur ces algorithmes pour la détection des troubles de santé. Cependant les troubles de santé les plus graves, affectant l'état général des animaux, étaient détectés avec une sensibilité moyenne à élevée (> 60%, pour une spécificité de 90%).
- En cas de détection, les alertes de santé générées par les algorithmes développés étaient plutôt précoces par rapport à l'évènement de santé à détecter (détection par

l'éleveur ou baisse de production laitière). Cela pourrait permettre dans certains cas de traiter de manière précoce des troubles de santé, même si la conduite à tenir pour l'éleveur lors d'alerte non associée à des signes cliniques pose question. Au vu des résultats, nous pouvons conseiller à l'éleveur de maintenir la surveillance des animaux en alerte durant quelques jours après la première alerte de santé.

- En présence d'un système de monitoring comme Heatime-HR®, susceptible de détecter une large variété de troubles de santé, plusieurs méthodes de référence peuvent être utilisées. Nous avons utilisé successivement 3 méthodes de référence : la détection par observation visuelle par l'éleveur, le repérage de chutes majeures de production laitière, et un examen clinique vétérinaire des vaches. Pour chacune des 3 méthodes de référence, une sensibilité faible à moyenne a été obtenue lorsque la spécificité des alertes de santé était élevée (> 80%). Ces 3 méthodes sont complémentaires. Elles permettent d'évaluer les alertes vis-à-vis de la détection d'évènements de nature différente.
- Activité et rumination sont complémentaires pour la détection des troubles de santé. Les performances de détection d'un système d'alerte combinant les algorithmes sur l'activité et la rumination étaient meilleures que celles des algorithmes sur l'activité ou la rumination considérées séparément.

En complément des alertes fournies par l'appareil Heatime-HR® pour la détection des chaleurs, dont la fiabilité a été démontré dans d'autres études et qui constitue l'utilisation principale du Heatime-HR®, des alertes de santé générées à partir des données d'activité et de rumination peuvent donc fournir des informations intéressantes aux éleveurs pour la détection des troubles de santé.

Cependant, une amélioration des performances des systèmes de détection des troubles de santé en général doit être opérée pour permettre une utilisation optimale, réelle et durable des informations fournies par ces systèmes. Plusieurs leviers d'amélioration sont identifiés comme la création de nouveaux capteurs, l'utilisation de nouvelles méthodes statistiques d'analyse des données, ou la combinaison d'informations provenant de différents capteurs ou systèmes d'information. L'expérience et les réflexions issues de cette thèse devraient être valorisées par la proposition de protocoles standardisés pour l'évaluation des systèmes de

monitoring de la santé, particulièrement lorsque ces systèmes sont susceptibles de détecter plusieurs types de maladies. Enfin, les technologies de l'élevage de précision allant très probablement être de plus en plus présentes dans les exploitations dans les prochaines années, un accompagnement du processus de changement par les organismes de développement agricole, les fabricants et vendeurs de ces technologies, ainsi que par les organismes de recherche est souhaitable. Ces technologies vont en effet probablement changer profondément le rapport des éleveurs à leur travail ainsi que la relation homme-animal.

Liste des publications et communications réalisées et prévues

Articles dans des périodiques à comité de lecture répertoriés dans BD internationales (ACL)

Acceptés

Clément, P., Guatteo, R., Delaby, L., Rouillé B., Chanvallon, A., Philipot, J.M., Bareille, N. 2014. Short communication: Added value of rumination time for the prediction of dry matter intake in lactating dairy cows. J Dairy Sci. DOI : 10.3168/jds.2013-7860.

Chanvallon A, Coyral-Castel S, Gatien J, Lamy JM, Ribaud D, Allain C, Clément P, Salvetti P. Comparison of three devices for the automated detection of estrus in dairy cows. Theriogenology. 2014 Sep 15;82(5):734-41. doi: 10.1016/j.theriogenology.2014.06.010.

Soumis

Clément, P., Guatteo, R., Quiniou, R., Bareille, N. Interest of monitoring drops in rumination time and neck activity for the detection of health disorders in dairy cows. J Dairy Sci, Soumis, 2014

En préparation

Clément, P., Guatteo, R., Bareille, N. Interest of monitoring drops in rumination time or in neck activity for the detection of milk production losses in dairy cows. Prev Vet Med, En preparation, 2014.

Communications par affiche dans un congrès international ou national (AFF)

Clément, P., Guatteo, R., Philipot, J.M., Lamy, J.M., Chanvallon, A., Trou, G., Bareille, N. 2013. Potential use of drops in rumination time for real-time detection of health disorders in dairy cows. In, Precision dairy conference, Rochester, Minnesota, USA, 26-27 juin.

Communications orales sans actes dans un congrès international ou national (COM)

Clément, P., Philipot, J.M.. 2012. An innovative tool for monitoring and phenotyping in cattle : Heatime-Heatime-HR®®. In, 38. International Committee for Animal Recording (ICAR), Cork, Ireland, 27 mai-1er juin.

Allain, C., Chanvallon, A., Clement, P., Bareille, N., Guatteo, R. Elevage de précision : périmètre, applications et perspectives en élevage bovin, In 21. Rencontres Recherches Ruminants (RRR), Paris, France. Accepté.

Clement, P., Guatteo, R., Bareille, N. 2012. Cost-benefit analysis of a putative earlier detection of health disorders in dairy cows based on a cow on-board monitoring tool: Heatime-HR®®. In, Proceeding Workshop in Animal Health Economics, Foulum, Danemark.

Conférences données à l'invitation du comité d'organisation dans un congrès national ou international (INV)

Clément, P., Guatteo, R., Bareille, N. 2014. Détection des troubles de santé par Heatime-HR®. In, Journées Nationales des Groupements Techniques Vétérinaires, Reims, France.

Annexe générale

Elevage de précision : périmètre, applications et perspectives en élevage bovin

C. ALLAIN¹, A. CHANVALLON², P. CLEMENT^{3,4}, R. GUATTEO^{3,4} et N. BAREILLE^{3,4}

¹Institut de l'Élevage, BP 85225, F-35652 Le Rheu Cedex

²Institut de l'Élevage, Oniris, F-44307 Nantes, France

³LUNAM Université, Oniris, UMR BioEpAR, CS 40706, 44307 Nantes, France

⁴INRA, UMR1300 BioEpAR, CS 40706, 44307 Nantes, France

RESUME

L'élevage de précision émerge depuis quelques années, notamment dans les exploitations bovines. Son développement s'explique par la disponibilité de nouvelles technologies dans les domaines de la microélectronique, de l'informatique, des télécommunications et des nanotechnologies, mais aussi par le besoin des éleveurs d'accroître la productivité et l'efficacité de leurs élevages. L'élevage de précision est principalement utilisé en élevage bovin laitier pour le pilotage de l'alimentation, la détection des troubles infectieux et des troubles métaboliques ou la surveillance des événements liés à la reproduction. Les performances de détection des outils commercialisés sont hétérogènes et très dépendantes des événements ciblés. Mais du point de vue des éleveurs, les performances et l'efficacité d'un outil est surtout caractérisé par son rapport coût/bénéfice. Les impacts économiques sont ainsi uniquement positifs pour des troupeaux de grandes tailles ou très productifs, mais de façon modérée. En revanche l'impact de l'élevage de précision sur l'évolution du métier d'éleveur, la relation homme/animal et l'organisation du travail ne sont pas négligeables et restent à mieux appréhender. Le développement des capteurs permet également d'envisager du phénotypage automatisé, standardisé et à haut débit sur des caractères non disponibles à ce jour et ouvre de nouvelles possibilités à la sélection génétique. Enfin, pour se développer davantage, l'élevage de précision doit offrir de nouvelles fonctionnalités aux éleveurs en ne s'arrêtant pas uniquement à la délivrance d'une information, mais en allant jusqu'à la fourniture d'un conseil.

SUMMARY

Precision livestock farming (PLF) has been emerging in the last few years, especially in dairy and beef farms. The growth of PLF is mostly pushed by the availability of new technologies in microelectronics, computer sciences, telecommunications, and nanotechnologies. But it also meets the farmers' needs to increase efficiency and productivity of their farms. PLF is mostly used in dairy farms to manage feeding, animal health and reproduction. Performances of PLF systems are very heterogeneous and closely dependent of the event targeted. But from a farmer point of view, the cost/benefit ratio is the main indicator of the PLF system performance and efficiency. Thus, it seems that economic impacts are only slightly positive and mostly for big or productive herds. On the other hand, impacts of PLF on farmers profession evolution, workload or man to animal relationship or not negligible and need to be better assessed. PLF development also opens new possibilities for standardized and automated phenotyping and selection on new characters. Finally, PLF needs to offer new features to improve its adoption amongst farmers, not only by supplying an information but also by providing an advice.

INTRODUCTION

Le développement de nouvelles technologies et leur utilisation en élevage, donnant naissance à l'élevage de précision, apparaît comme un des leviers d'action possible pour répondre au besoin de développement d'un élevage durable économiquement, socialement et respectueux de l'environnement (Eastwood *et al.*, 2004 ; Bewley, 2010). Plusieurs raisons expliquent le développement de l'élevage de précision au cours de ces dernières années. Quel que soit le type de production ou l'espèce concernés, la taille des élevages français et européens augmente chaque année (Gambino *et al.*, 2012) sans augmentation de la main d'œuvre associée. Cette évolution est en partie motivée par les perspectives d'évolution des politiques agricoles communes (fin des quotas), la volatilité du coût des matières premières et du prix de vente des produits qui rendent les marges économiques plus minces que par le passé

(Bewley, 2010). Par conséquent, l'amélioration, même faible, de la productivité ou de l'efficacité des élevages est devenue un enjeu important pour améliorer leur profit. De plus, les éleveurs doivent faire face aux demandes sociétales et aux contraintes législatives croissantes autour du bien-être animal, de l'impact de l'élevage sur l'environnement ou encore de la rationalisation de l'utilisation d'intrants médicamenteux (Berckmans, 2004, Wathes, 2007). Enfin, les éleveurs aspirent de plus en plus à une diminution de leurs contraintes (traite, distribution de l'alimentation, observation des chaleurs,...) et à une meilleure maîtrise du temps et du confort de travail. Parallèlement à ces constats, des sauts technologiques dans la microélectronique, l'informatique, les télécommunications et désormais les nanotechnologies, ont eu lieu. Leur utilisation dans d'autres industries a permis de réduire

considérablement leur coût. Ainsi, par exemple, le succès mondial des téléphones mobiles ou encore des jeux vidéo a eu un impact important sur la diminution des coûts des technologies de communication sans fil ou des capteurs. Cela favorise de fait leur utilisation pour d'autres applications comme l'élevage de précision. Néanmoins, passé l'effet de mode, il convient d'identifier les conditions optimales de mise en œuvre et de déploiement de telles technologies en élevage.

Cette synthèse reprend l'état des connaissances sur l'élevage de précision et fait le point sur les perspectives permises pour l'élevage bovin laitier et allaitant.

1. DEFINITION DE L'ELEVAGE DE PRECISION

Différentes définitions de l'élevage de précision sont proposées dans la littérature. Bewley (2010) définit l'élevage laitier de précision comme l'utilisation de technologies permettant de mesurer des indicateurs physiologiques, comportementaux ou de production sur les animaux pour améliorer les stratégies de management du troupeau et les performances de l'élevage. Ces performances peuvent être économiques, sociales ou environnementales (Eastwood *et al.*, 2004, Bewley, 2010). Selon Berckmans (2012), c'est le pilotage de l'élevage grâce au monitoring automatisé et en temps réel de la production, de la reproduction, de la santé et du bien-être des animaux.

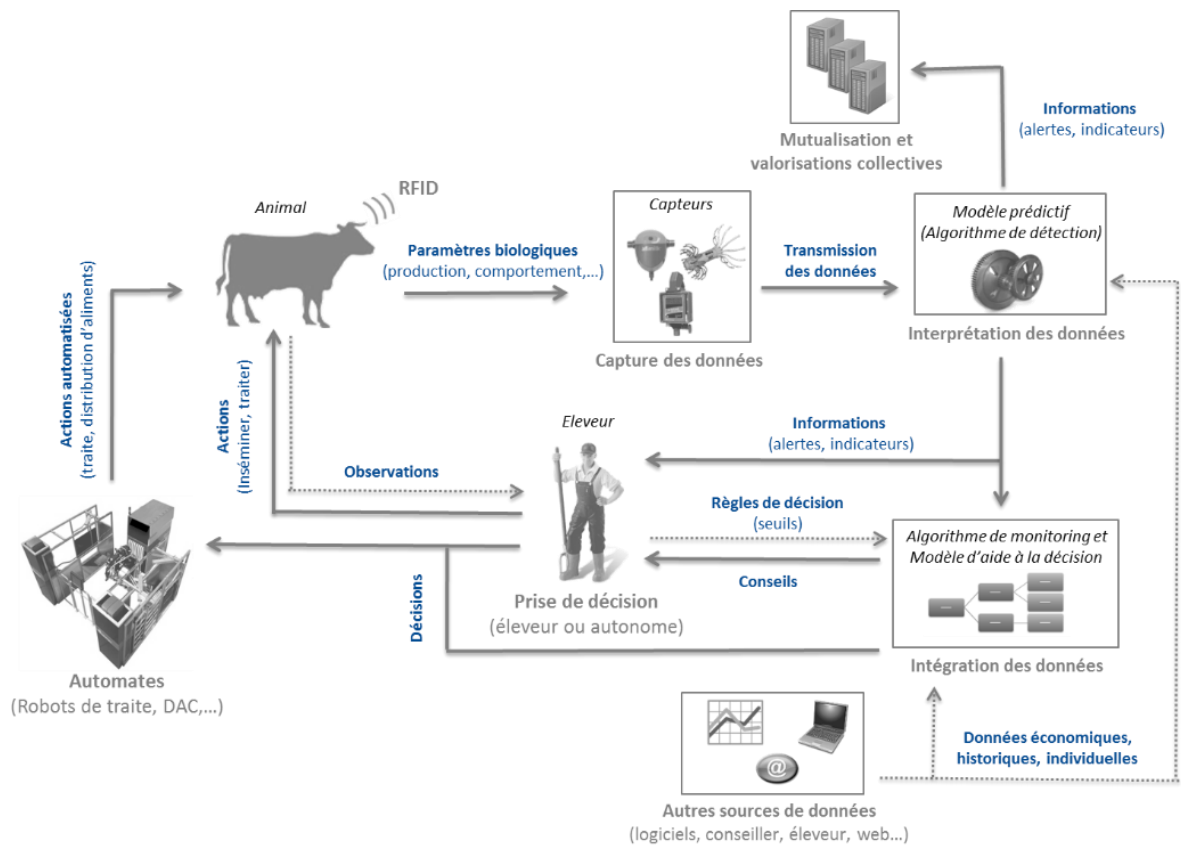
Ainsi, une définition relativement consensuelle aux différentes filières animales peut être proposée (Hostiou *et al.*, 2014) : l'élevage de précision, c'est l'utilisation coordonnée i) de capteurs pour mesurer des paramètres comportementaux, physiologiques ou de production sur les animaux, ii) de TIC (Technologies de l'Information et de la Communication) pour échanger, stocker, transformer et restituer ces informations à l'éleveur

pour iii) l'aider dans sa prise de décision en complément de ses observations. Par extension, l'utilisation d'automatismes permettant de décharger l'éleveur de certaines tâches astreignantes (traite, alimentation, régulation de l'ambiance des bâtiments) peut également être associée à l'élevage de précision s'ils sont couplés à l'utilisation de capteurs et de technologies de transfert d'informations. D'autres automatismes (robots racleurs des déjections, pousse-fourrage, « chiens » de troupeaux pour pousser les vaches en salle de traite, ...) ne produisant pas de données utilisables pour la conduite du troupeau sont exclus de la notion d'élevage de précision.

Selon Aerts *et al.* (2003), Berckmans (2004) et Whates (2007), plusieurs conditions doivent être réunies pour permettre un monitoring et un pilotage continu de l'élevage. D'abord, des variables animales doivent être mesurées et analysées en continu à une échelle de temps appropriée à l'aide de capteurs fiables. Dans un second temps, un modèle prédictif (algorithme de détection) pour chaque variable mesurée (ou combinaison de variables) de la réponse de l'animal aux conditions environnementales (alimentation, climat, conduite d'élevage...) doit être développé. C'est ensuite la comparaison entre ce qui est attendu (calculé par ce modèle) et ce qui est mesuré par les capteurs qui va permettre de détecter les animaux ou les situations (CEstrusœstrus par exemple) nécessitant une attention particulière de la part de l'éleveur. Une information sous forme d'alerte ou d'indicateur lui sera envoyée.

Un niveau supplémentaire d'intégration de cette information avec d'autres données (économiques par exemple) peut permettre de produire un conseil à l'éleveur (Rutten *et al.*, 2013). Ce niveau sera décrit plus précisément dans la partie 4.3. Le concept d'élevage de précision est représenté en Figure 1.

Figure 1 : Concept d'élevage de précision (adapté de Allain *et al.*, 2012 ; Rutten *et al.*, 2013)



2.1 QUELS OUTILS POUR QUELLES MESURES ?

Dans le cadre de l'élevage de précision, les capteurs permettent (1) d'automatiser une mesure pouvant être réalisée manuellement (pesée de la quantité d'aliment ingérée) ou visuellement (chevauchements, vêlages) ; ou (2) de mesurer de nouveaux paramètres non détectables par l'œil humain ou difficilement mesurables (activité, rumination). Ces capteurs peuvent être positionnés sur l'animal, intégrés dans un automate de traite, de pesée, de tri ou d'alimentation ou positionnés dans les bâtiments d'élevage. D'après Faverdin (2012), ces outils basés sur des technologies diverses peuvent en général mesurer des paramètres :

- physiologiques sur l'animal (production laitière, ingestion, température corporelle),
- biologiques sur ses produits (caractéristiques physico-chimiques du lait, niveaux hormonaux),
- morphologiques (mensurations, poids)
- comportementaux (déplacement, activité, comportement alimentaire).

Les données produites sont ensuite valorisées soit pour le management du troupeau (alimentation), soit pour la détection précoce de troubles de la santé (mammites troubles métaboliques) ou d'évènements d'intérêt (vêlages, chaleurs). Dans certains cas, les informations mesurées par ces capteurs et croisées avec d'autres informations peuvent également servir à modifier le fonctionnement d'un automate (fréquence de traite, quantité de concentrés distribuée, séparation du lait ou d'un animal). Un panorama des principaux outils disponibles est présenté dans le tableau 1.

En complément, de nombreux capteurs sont en phase de développement ou d'adaptation à l'élevage bovin. Des innovations sont à l'étude pour le pilotage fin du troupeau (note d'état corporel, spectrométrie moyen infra rouge) et seront développées dans la suite du document. Dans le domaine de la santé, d'autres applications peuvent être citées : la détection de troubles respiratoires chez les porcs par l'enregistrement des sons liés à la toux qui pourrait être appliquée aux bovins notamment aux veaux (Ferrari *et al.*, 2010) et la détection de boiteries par imagerie numérique (Viazzi *et al.*, 2013). Enfin le bâtiment d'élevage fait aussi l'objet d'avancées technologiques par exemple concernant la mesure de consommation d'eau (Ménard *et al.*, 2012) ou le renouvellement d'air pour préserver la santé des animaux. Des capteurs sont ainsi mis au point pour réguler de façon automatique plusieurs paramètres d'ambiance : température, hygrométrie, concentrations en ammoniac et en dioxyde de carbone (Fang, 2003 ; Janni et Jacobson, 2013).

2.2 PERFORMANCES POUR LA DETECTION D'EVENEMENTS SANITAIRES ET DE REPRODUCTION

L'évaluation des performances des outils de monitoring de la santé et de la reproduction est capitale pour permettre (i) un choix éclairé des éleveurs lors de l'investissement et (ii) un usage réel, conforme et durable de ces outils en élevage. En effet, pour la santé par exemple, un système générant trop d'alertes chez des animaux sains

(faux positifs), ou pas suffisamment d'alertes chez des animaux malades (faux négatifs), risque d'entraîner une perte de confiance de l'éleveur dans les informations fournies par le système et une sous-utilisation de celui-ci. Mal utilisé, il peut engendrer des atteintes au bien-être animal, une surutilisation des traitements médicamenteux et impacter négativement l'économie des exploitations. L'évaluation des performances d'un système de détection se fait par le calcul de statistiques comme le pourcentage d'évènements à détecter (chaleurs ou troubles de santé) effectivement détectés par le système (sensibilité), le pourcentage d'évènements à ne pas détecter (absence de chaleurs ou de troubles de santé) effectivement non détectés par le système (spécificité) et l'évaluation de la précocité des détections. La sensibilité et la spécificité constituent les valeurs intrinsèques d'un système d'alerte. Pour l'éleveur, il s'agit plutôt de savoir quelles seront les performances d'un système donné dans son élevage. Pour cela, on utilise des statistiques opérationnelles comme la valeur prédictive positive (VPP=probabilité qu'une vache en alerte ait l'évènement à détecter, chaleur ou trouble de santé) ou négative, ou le taux de détection à tort (par exemple le nombre de fausses alertes pour 1000 traites pour un système de détection des mammites) qui permettent de rendre compte simplement de la qualité du système d'alertes (Kamphuis *et al.*, 2008). Ces valeurs prédictives des alertes dépendent non seulement de leurs sensibilité et spécificité, mais aussi de la prévalence des évènements à détecter dans l'élevage : ainsi plus la prévalence est forte, plus la VPP sera élevée, ce qui sera recherché en priorité en élevage (détection des vaches réellement malades).

Or, la prévalence des troubles de santé en élevage est heureusement relativement faible. Pour atteindre des valeurs prédictives qui soient suffisamment élevées et que les outils de monitoring apportent réellement une information utile à l'éleveur, il faut ainsi que la sensibilité et la spécificité des alertes soient très élevées. Ainsi, la norme ISO/FDIS 20966 (ISO, 2007) recommande une sensibilité supérieure à 80% combinée à une spécificité supérieure à 99% pour les systèmes de détection automatisée des mammites sur les robots de traite, ce qui permet d'atteindre des valeurs de VPP de 35-50%. Cependant, parmi les 28 études citées par Rutten *et al.* (2013) et ayant évalué récemment des systèmes automatisés de détection des mammites, seules 3 systèmes d'alerte atteignent ces objectifs. C'est alors à l'éleveur de « filtrer » parmi les nombreuses alertes générées par le système sur des animaux sains (défaut de spécificité) celles qui requièrent réellement son attention. Il peut pour cela utiliser les autres informations dont il dispose (historique de la vache, observation visuelle, autres capteurs...). Cependant, la pertinence du tri effectué est subjective et l'objectif de diminuer la charge mentale et physique associée à la gestion d'une maladie n'est dans ce cas pas atteint (Dreno, 2009). Parmi les technologies fournissant des alertes de santé, la plupart sont donc moins performantes que l'éleveur pourtant déjà peu sensible (exemple des boiteries,

Fabian *et al.*, 2014), ce qui devrait être le niveau d'exigence minimal pour leur utilisation en élevage. Au contraire, les systèmes automatisés de détection des chaleurs atteignent des sensibilités de 50 à 99% pour des spécificités supérieures à 93% (Rutten *et al.*, 2013), ce qui correspond à des performances généralement supérieures aux performances moyennes de détection des éleveurs autour de 50% (At-Taras et Spahr, 2001). Des systèmes de détection de l'ovulation (dosage de la progestérone dans le lait) sont désormais disponibles sur le marché et sont encore plus performants (sensibilité de 93 à 99%, spécificité de 94% ; Friggens *et al.*, 2008) (Tableau 1).

L'évaluation technique des outils doit être regardée de manière critique car les méthodes d'évaluation utilisées ne sont pas standardisées. Une problématique majeure réside dans la méthode de référence à laquelle l'outil est confronté, le gold standard parfait n'existant souvent pas. On se retrouve ainsi fréquemment à comparer l'outil à l'observation de l'éleveur que l'on sait déjà par ailleurs subjective. Par exemple, la méthode de référence utilisée pour classer les animaux sains ou malades peut être variable d'une étude à l'autre (concentration en cellules somatique ou observation visuelle pour les mammites) ou peut être très imparfaite. Ainsi, dans une étude récente (Timsit *et al.*, 2011), seuls 26% des hyperthermies détectées chez des broutards par un thermobolus avaient été détectées et traitées par les éleveurs. La détection par l'éleveur aurait donc été une méthode de référence très imparfaite pour évaluer ce thermobolus.

Cependant, si la connaissance des performances des outils est importante, l'évaluation doit aussi intégrer l'impact économique de l'équipement et l'impact sur l'organisation du travail de l'éleveur.

2. CONSIDERATIONS ECONOMIQUES ET SOCIALES

3.1 IMPACTS TECHNIQUES ET ECONOMIQUES

Pour un éleveur, le coût des outils de monitoring est assez facile à cerner. Par exemple, pour les outils de détection des chaleurs, il comprend l'investissement (base réceptrice et un set de capteurs) et la maintenance (changement des piles, casse). Ainsi, pendant les 5 premières années d'utilisation, le coût par vache et par an varie de 20 € dans les grands troupeaux à 60 € dans les petits troupeaux et dans ceux à vêlages étalés qui nécessitent d'équiper toutes les vaches (Courties, 2014).

En revanche, les bénéfices économiques (hors temps de travail) sont plus difficilement évaluables. Ils vont essentiellement dépendre de l'amélioration de la sensibilité de détection des événements obtenue grâce à l'équipement en capteur des animaux. Ainsi, une meilleure détection des vêlages doit améliorer la survie des veaux ; une meilleure détection des vaches malades doit permettre de limiter leurs pertes de production, mais, si un traitement médicamenteux avec temps d'attente est appliqué, ce bénéfice peut être vite perdu.

Les travaux de quantification des bénéfices technico-économiques attendus d'un équipement de surveillance ont surtout concerné les détecteurs de

chaleurs. Ils peuvent permettre de raccourcir le délai de mise à la reproduction, les intervalles entre inséminations non-fécondantes, l'intervalle entre vêlages et in fine d'améliorer la marge brute de l'atelier laitier (Seegers *et al.*, 2010). Rutten *et al.* (2013) ont analysé l'intérêt technico-économique de l'équipement en détection automatisée des chaleurs dans un troupeau de 130 vaches laitières par modélisation stochastique. Ils ont montré qu'avec une amélioration de 30 % de sensibilité de détection, il était possible d'atteindre un retour sur investissement en 7 ans. Des résultats similaires ont été observés par Østergaard *et al.* (2005) et Dolecheck *et al.* (2013) en grands troupeaux. Une exploration récente dans le contexte français (Courties, 2014) confirme en partie ces conclusions. En effet, le coût de l'équipement n'est compensé en 5 ans que pour les troupeaux de grande taille (120 vaches laitières et plus) et pour ceux de taille moyenne (60 vaches) à fort niveau de production, là encore si l'amélioration de sensibilité atteint 30 à 40%. Ces niveaux d'amélioration peuvent être attendus dans des exploitations où le temps consacré à l'observation des chaleurs avant équipement est très limitant. A contrario, pour des éleveurs performants avant équipement (sensibilité >50%) et pour les troupeaux à production modérée ou de taille réduite, l'investissement est difficilement rentable économiquement.

L'évaluation du rapport coût/bénéfices des nouvelles technologies en élevage est donc un enjeu important pour leur développement. Des travaux en cours visent à développer un outil d'aide à la décision pour s'équiper d'un capteur pour la détection automatisée des chaleurs, incluant une analyse du rapport coût/bénéfices.

3.2 IMPACTS SUR L'ORGANISATION DU TRAVAIL

Une synthèse récente proposée par Hostiou *et al.* (2014) dresse un état des lieux complet de l'impact de l'élevage de précision sur l'organisation du travail et le métier d'éleveur. En effet, les automatismes (robots de traite ou d'alimentation), couplés à la collecte de données par des capteurs et leur restitution sous forme d'alertes ou de rapports synthétiques, peuvent améliorer le confort de travail en réalisant automatiquement des tâches pénibles, en aidant à la prise de décision ou en allégeant la charge mentale par l'indication des interventions nécessaires, par exemple, à l'occasion des mises-bas. En revanche, de nouvelles tâches peuvent en découler comme l'apprentissage de son utilisation, l'entretien et la surveillance de l'équipement, la consultation et l'interprétation des données. Cela peut être source de stress en raison de la multiplicité des alarmes ou des alertes, parfois la nuit, et des risques accrus de pannes dans des systèmes électroniques sophistiqués. La complexité des informations à gérer peut désorienter les éleveurs et être à l'origine d'erreurs d'interprétation et de diagnostics. Enfin, l'élevage de précision peut conduire à un réel désinvestissement des exploitants dans les savoir-faire pratiques (repérage visuel des animaux malades ou en chaleur,...). L'élevage de précision est susceptible d'exercer un attrait pour le métier notamment de jeunes en quête

de modernité, mais peut aussi se révéler source d'échecs s'il n'est pas adapté aux besoins et aux compétences des éleveurs. Un article détaillant ce thème est proposé lors de ces journées 3R 2014.

3. LES PERSPECTIVES DE L'ELEVAGE DE PRECISION POUR LE PHENOTYPAGE HAUT DEBIT ET LA CONDUITE DU TROUPEAU

4.1 DE NOUVELLES POSSIBILITES POUR LA MESURE DE PHENOTYPES D'INTERET EN ELEVAGE

En permettant d'évaluer des paramètres non observables par l'homme ou difficilement mesurables, l'élevage de précision apporte de réelles perspectives pour mesurer des phénotypes d'intérêt.

L'estimation de l'état des réserves corporelles est un bon exemple d'application potentielle (Boichard et Brochard, 2012). Celles-ci sont en effet fortement liées à la productivité, la longévité ou au bien-être des vaches. De plus, les variations des réserves corporelles reflètent le bilan énergétique de l'animal qui constitue un indicateur intéressant pour le pilotage de troupeau. Mais il n'est mesuré que peu précisément dans les élevages expérimentaux (via la Note d'Etat Corporel, NEC) et rarement estimé dans les élevages commerciaux.

Récemment, Fischer *et al.* (2014) ont développé une méthode pour estimer la NEC à partir de l'utilisation de l'imagerie en 3D. La méthode construite à partir d'un traitement global de l'image a permis d'estimer objectivement la NEC des vaches laitières. La mesure de la NEC par les surfaces 3D apparaît surtout 3 fois plus répétable (écart-type = 0,07 versus 0,21 pour la NEC) et 2,8 fois plus reproductible que la NEC estimée par les notateurs (écart-type= 0,1 versus 0,28 pour la NEC).

Un autre exemple est la spectrométrie moyen infrarouge (MIR). Cette technique est utilisée depuis longtemps dans les laboratoires d'analyse du lait pour déterminer les éléments de paiement du lait : taux butyreux, taux protéique, comptage de cellules somatiques, mais aussi d'autres caractères comme l'urée ou le lactose. Depuis quelques années, l'utilisation de cette technologie s'est développée grâce à la possibilité d'extraire et de stocker les spectres qui, jusqu'alors, étaient détruits après détermination des caractères souhaités. De nombreux projets ont vu le jour pour fournir aux éleveurs des indicateurs de pilotage du troupeau fondés sur l'analyse des données spectrales. Les plus aboutis concernent la composition en acides gras et protéines du lait pour le pilotage fin de l'alimentation et de l'indexation génomique (Gelé *et al.*, 2014 ; Ferrand-Calmels *et al.*, 2014 ; De Marchi *et al.*, 2013). Depuis, les recherches sur les spectres MIR se sont étendues à d'autres caractères quantitatifs (teneurs en minéraux et en corps cétoniques du lait) et qualitatifs (bilan énergétique, maladies métaboliques, aptitude fromagère du lait, diagnostic de gestation, etc) (De Marchi *et al.*; 2014). Ces applications font toujours l'objet de projets de recherche et développement. La spectrométrie MIR présente donc de nombreux

avantages – rapidité, faible coût – qui permettraient à l'avenir une utilisation en routine et à grande échelle de nombreux phénotypes pour le pilotage fin du troupeau (équilibre des rations, couverture des besoins, fertilité, santé, empreinte environnementale) mais aussi pour la constitution de population de référence en vue d'une indexation génomique sur ces nouveaux caractères (De Marchi *et al.*, 2014).

En revanche, certaines autres potentialités prometteuses ont été également évaluées avec des résultats malheureusement plus mitigés. Ainsi, une étude récente s'est intéressée à la relation entre durées de rumination et quantités ingérées chez des vaches laitières en lactation. La mesure des durées de rumination avec une bonne précision est en effet possible depuis 2011 en élevage commercial, grâce à un appareil de monitoring de la santé (Schirmann *et al.*, 2009). L'hypothèse faite par les auteurs était que les durées de rumination étaient liées aux quantités ingérées, une durée de rumination plus importante étant nécessaire pour remastiquer une quantité croissante de ration. Les durées de rumination auraient ainsi pu permettre d'évaluer indirectement les quantités ingérées et donc d'améliorer les modèles de prédiction de l'ingestion voire de contribuer à estimer l'efficacité alimentaire des animaux. L'étude a cependant montré qu'il n'existait pas de relation entre les durées de rumination et les quantités de matière sèche ingérées durant la lactation chez des vaches recevant une ration complète (Clément *et al.*, 2014).

4.2 VERS UNE UTILISATION POUR LE PHENOTYPAGE INDIVIDUEL DE MASSE

En France, près de 4000 exploitations laitières sont aujourd'hui équipées de robots de traite (Idele, 2014) avec leurs lots de capteurs qui en font de véritables plateformes potentielles de phénotypage.

Ce développement très rapide (nombre d'élevages équipés multiplié par 4 depuis 2007) ouvre la voie à de nouvelles perspectives pour le phénotypage animal à grande échelle sur de nouveaux caractères ou des caractères existants mais mesurés avec une faible précision ou une fréquence inadaptée. Ainsi, des phénotypes tels que la vitesse de traite (Heringstad, 2014), les débits de matière utile (Fogh, 2012), le tempérament (basé sur le comportement pendant la traite mesuré par le nombre de fois ou les faisceaux trayeurs se décrochent ; Rinell, 2013), la conformation de la mamelle (basée sur les enregistrements des caméras des robots ; Fogh, 2012) pourraient facilement être utilisés pour la sélection. Les données de conductivité du lait pourraient également être utilisées pour le phénotypage de la santé mammaire (Norberg, 2005).

De nombreux robots de traite sont également équipés de balances électroniques qui pèsent les animaux lors de chaque traite. La pesée automatique plusieurs fois par jour est un moyen prometteur de prévoir les changements de poids quotidiens, l'accumulation des mesures permettant de contrôler les variations aléatoires dues à l'alimentation et l'abreuvement (Boichard et Brochard, 2012).

De nouveaux dispositifs comme le Herd Navigator réalisent une prise d'échantillon et une analyse en ferme de marqueurs biologiques comme la Lactate Deshydrogénase pour la détection des mammites, le beta hydroxy butyrate pour les acétonémies ou la progestérone pour la reproduction et permettent des diagnostics plus précis (Friggens *et al.*, 2008 ; Hojsgaard et Friggens, 2010).

Les capteurs embarqués sur les animaux sont aussi des sources prometteuses de données pour le phénotypage. Les systèmes de détection automatisée des chaleurs comme les podomètres et les accéléromètres sont en plein essor. Outre la production d'alertes, ils pourraient également être utilisés pour caractériser l'expression des chaleurs et la cyclicité des vaches laitières (Lovendhal and Chagunda, 2009). Ces capteurs permettent également de disposer de données en continu sur des grandeurs physiologiques parfois méconnues permettant de mieux comprendre la physiopathologie de certains troubles.

Les accéléromètres positionnés sur le cou permettent de plus en plus d'avoir accès au comportement alimentaire (ingestion, rumination, abreuvement) en plus de la détection des Œstrus. Ils pourraient être un bon complément des mesures réalisées par les DAC (Distributeur Automatique de Concentrés) pour les travaux sur l'efficacité alimentaire.

De nombreux autres capteurs (température et pH ruminal, etc. – Tableau 1) peuvent être envisagés pour des applications quasiment illimitées, comme par exemple l'identification d'animaux résistants ou résilients à certaines maladies/affections.

Ces technologies offrent donc des perspectives de phénotypage à grande échelle (plusieurs dizaine voire centaine de milliers d'animaux). L'avantage principal est que les coûts de collecte de ces informations sont limités et que les données sont enregistrées de façon objective et avec une fréquence élevée, à condition que les appareils soient correctement réglés.

En revanche, un challenge très important concerne la standardisation de ces données et l'accès à celles-ci (y compris leur propriété et leur droit d'utilisation potentielle), sachant que chaque constructeur fonctionne avec son propre format de données et ses propres logiciels (Boichard et Brochard, 2012). Des travaux sur la standardisation des échanges de données entre les robots, les capteurs et les organismes d'élevage sont d'ailleurs en cours au sein de l'International Committee for Animal Recording (ICAR) (Rehben, 2014).

4.3 VERS UNE PRISE DE DECISION OPTIMISEE OU AUTONIMISEE

La prise de décision (et l'action qui en découle) est l'étape ultime et indispensable de tout processus de management de l'exploitation. En élevage, certaines décisions comme l'application d'un traitement médicamenteux ou l'insémination peuvent impacter négativement l'économie de l'exploitation, la santé ou le bien-être animal si elles sont prises à tort ou si elles ne sont pas prises au bon moment. Les technologies d'élevage de précision, en apportant une information objective, et si possible fiable, à l'éleveur, le sécurisent dans sa prise de décision.

Récemment, Rutten *et al.* (2013) ont décrit la dynamique de l'information inhérente à l'élevage de précision, du capteur à la décision (Figure 1). Dans cette publication, les auteurs soulignent que les travaux scientifiques concernant les outils d'élevage de précision n'évaluent souvent que leurs performances pour détecter des troubles de santé ou des événements de reproduction. Pourtant, l'agrégation d'informations issues des capteurs avec d'autres informations (économiques, historiques, techniques) provenant de l'élevage pourraient permettre la production de conseils pour faciliter la décision de l'éleveur et nécessite d'être creusé.

Cette intégration de modèle d'aide à la décision par les constructeurs représente une perspective d'avenir potentiellement intéressante pour l'élevage de précision. Quelques dispositifs commerciaux le proposent déjà.

Ainsi, le système DLM® (Dynamic Linear Model) développé par Lely en lien avec ses robots de traite, permet d'ajuster automatiquement la distribution de concentrés et la fréquence de traite individuellement en fonction de considérations techniques (niveau de production, TB et TP du lait, intervalle de traite et temps de préparation à la traite) et économiques (prix du lait et coût des aliments) pour optimiser la marge économique par litre de lait et non plus uniquement le niveau de production.

De même, le système Herd Navigator® (DeLaval + Foss), prend en compte des facteurs comme la parité de la vache, son stade de lactation, ou son niveau de production en complément de l'alerte issue du dosage de progestérone pour conseiller à l'éleveur d'inséminer ou non.

Dans ces deux exemples, on pourrait craindre que l'éleveur perde la maîtrise de sa décision voire de l'action qui en découle. En réalité, il peut rester maître des règles de décision qui sont appliquées (par exemple le stade de lactation en dessous duquel il ne souhaite pas inséminer l'animal) et délègue la décision et parfois l'action à un automate (distribution de concentrés, par exemple).

4.4 DE NOUVELLES ECHELLES D'APPLICATION

Enfin, à la problématique de la prise de décision s'ajoute celle de l'échelle d'application. Si le monitoring en bovins est majoritairement individualisé, les données individuelles peuvent être agrégées pour produire de l'information nouvelle. Rutten *et al.* (2013) détaillent ainsi deux types de valorisation des données à l'échelle du troupeau. D'abord, ces données peuvent être réutilisées pour améliorer l'algorithme de détection. Par exemple, dans le cas de la détection des chaleurs basée sur l'activité physique, une hausse générale d'activité du troupeau liée à un événement particulier (sortie au pâturage) peut être utilisée pour corriger l'algorithme de détection individuel. Ensuite, ces données « troupeau » peuvent apporter une information nouvelle à l'éleveur sur l'état de son cheptel. Ainsi, une baisse de l'intensité d'expression des chaleurs ou la diminution de la fréquence des alertes pourrait être un signe d'une détérioration de l'état reproductif du troupeau avant que l'éleveur s'en aperçoive par lui-même. Enfin, une agrégation des données à une échelle collective plus importante (non plus à l'échelle de la ferme, mais à

une échelle régionale ou nationale) pourrait servir à la sélection sur de nouveaux caractères (cf. 4.3), mais aussi à détecter des dérives de performances concomitantes à l'arrivée d'un évènement sanitaire ou climatique. Ce niveau d'information permettrait alors d'évaluer l'impact d'une maladie émergente plus précocement qu'aujourd'hui (exemple de la FCO) et de proposer un plan de maîtrise rapidement.

CONCLUSION

Cet article présente la diversité d'applications de l'élevage de précision en élevage bovin pour le monitoring de la santé et de la reproduction. Il souligne également que si les outils d'élevage de précision semblent être d'un intérêt évident pour le gain de temps et de confort de travail, ils ne sont pas toujours rentables techniquement et surtout économiquement. Des méthodes d'évaluation du rapport coût/bénéfices de ces technologies restent cependant à construire pour mieux appréhender ces premières constatations.

Si l'élevage de précision se développe, il est évident qu'il en est encore à un stade précoce. D'abord, les fonctionnalités offertes à l'éleveur doivent franchir le cap de la délivrance d'une alerte (œstrus, mammite) pour aller jusqu'à la fourniture d'un conseil (insémination, traitement). Ensuite, le développement technologique actuel offre de nouvelles opportunités pour approcher des paramètres clés difficilement mesurables mais cruciaux pour l'avenir de la production bovine (efficacité alimentaire, bien-être, etc.). Enfin, les outils déjà déployés en masse sur le terrain pourraient être mieux valorisés pour le phénotypage et la sélection, mais aussi comme sentinelles pour suivre l'évolution d'épizooties ou de maladies émergentes.

BIBLIOGRAPHIE

Aerts, J.-M., Wathes, C.M., Berckmans, D., 2003. Biosyst Eng, 84, 257-266.
Allain, C., Duroy, S., Alix, E., Dassé, B., Delaunay, M., Langlais, J., 2012. Conf. SPACE 2012.
At-Taras, E.E., Spahr, S.L., 2001. J. Dairy Sci., 84, 792-798
Berckmans D., 2004. In International society for animal hygiene. 27-30.
Berckmans D., 2012.
http://ec.europa.eu/food/animal/diseases/strategy/docs/presentation_15062012_point_1_berckmans_en.pdf
Bewley, J., 2010. In: Proc. First North Am. Conf. Precis. Dairy Manag., Toronto, Canada, 16p
Bewley J.M., Russell R.A., 2010. First North Am. Conf. Precis. Dairy Manag., Toronto, Canada, 30-31
Boichard, D., Brochard M., 2012. Animal, 6:4, 544-550
Clément, P., Guatteo, R., Delaby, L., Rouillé, B., Chanvallon, A., Philipot, J.M, Bareille, N., 2014. J. Dairy Sci. <http://dx.doi.org/10.3168/jds.2013-7860>.
Dolecheck, K.A., Heersche, G., Bewley, J.M., 2013. Precision Dairy Conference, Minnesota, USA.
Courties R., 2014. Thèse DV, Oniris Nantes, 119 p.
De Marchi, M., Toffanin, V., Cassandro, M., Penasa, M., 2013. J. Dairy Sci., 96, 4707-4715

De Marchi, M., Toffanin, V., Cassandro, M., Penasa, M., 2014. J. Dairy Sci., 97, 1171-1186
Dreno, J.N., 2009. Thèse DV, Oniris Nantes, 134 p.
Eastwood C., Chapman D., Paine M., 2004. Precision dairy farming-taking the microscope to dairy farm management.
Fabian, J., Laven, R. A., Whay, H. R., 2014. Vet. J., 201, 31-38
Fang, W., 2003. http://www.agnet.org/library.php?func=view&id=20110801163637&type_id=5
Faverdin, P., 2012. EAAP Annual Meeting, Bratislava, Slovaquie.
Ferrand-Calmels et al., 2014. J. Dairy Sci., 97, 1-19
Ferrari, S., Piccinini, R., Silva, M., Exadaktylos, V., Berckmans, D., Guarino, M., 2010. Prev. Vet. Med., 96, 276-280
Fischer, A., Luginbuhl, T., Delattre, L., Delouard, J.M., Faverdin, P., 2014. Renc. Rech. Ruminants, 21
Fogh, A., 2012. 38th ICAR annual meeting, Cork (Ireland)
Friggens, N.C., Bjerring, M., Ridder, C., Hojsgaard, S., Larsen, T., 2008. Reprod Dom Anim Vol 43.
Gambino M., Laisney C., Vert J., 2012. http://agriculture.gouv.fr/IMG/pdf/Le_monde_agricole_en_tendances.pdf
Gelé, M. et al., 2014. Innovations Agronomiques, 34, 189-206
Herrigstad, B., 2014. 39th ICAR annual meeting, Berlin (Germany)
Hojsgaard, S., et Friggens, N.C., 2010. J. Dairy Sci., 93, 582-592.
Hostiou, N., Allain, C., Chauvat, S., Turlot, A., Pineau, C., Fagon, J., 2014. INRA Prod. Anim., 2014, 27 (2), 113-122
Institut de l'élevage, 2014. <http://idele.fr/domaines-techniques/sequiper-et-sorganiser/elevage-de-precision/publication/idelesolr/recommends/robots-de-traite-le-deploiement-continue.html>
ISO 20966:2007. Geneva, Switzerland.
Janni, K.A., Jacobson, L.D., 2013. Prec. Dairy Conf., Rochester, June 25 - 27, 2013, 63-64
Kamphuis, C., Sherlock, R., Jago, J., Mein, G., Hogeveen H., 2008. J. Dairy Sci., 91, 4560-4570
Lovendahl, P., Chagunda, M.G.G., 2009. J. Dairy Sci. 93 :249-259
Ménard, J.L. et al., 2012. Renc. Rech. Ruminants, 19, 173-176.
Norberg, E., 2005. Liv. Prod. Sci., 96, 2-3, 129-139
Ostegaard, S, Friggens, N.C, Chagunda, M.G.G., 2005. Theriogenology. Vol 64.
Rehben, E., 2014. 39th ICAR annual meeting, Berlin (Germany)
Rinell, E., 2013. Thèse Master, Upsalla, 24p
Rutten, C.J., Steeneveld, W., Inchaisri, C., Hogeveen, H., 2013. Precision livestock farming. Wageningen Academic Pub, 124-132.
Rutten, C.J., Velthuis, A.G.J, Steeneweld, W., Hogeveen, H., 2013. J. Dairy Sci. 96 :1928-1952
Schirmann, K., von Keyserlingk, M. A., Weary, D. M., Veira, D. M., Heuwieser, W., 2009. J. Dairy Sci., 92, 6052-6055
Timsit, E., Assie, S., Quiniou, R., Seegers, H., Bareille, N., 2011. Vet. J., 190, 136-142

- Seegers, H., Billon, D., Bossard-Apper, E., Ponsart, C., Bareille, N., 2010.** Renc. Rech. Ruminants, 17, 146.
- Timsit, E., N. Bareille, H. Seegers, A. Lehebel, and Assie. S., 2011a.** J Anim Sci 89:4272-4280.
- Viazzi, S., et al., 2013.** J. Dairy Sci., 96, 257-266
- Wathes C., 2007.** In XIIIth International Congress in Animal Hygiene. 397-404.

Tableau 1 : Panorama des principaux outils de monitoring disponibles en élevage bovin et leurs applications

Système	Mesure	Applications	Performances de détection ⁽¹⁾	Coût (50 vaches)
Paramètres physiologiques sur l'animal				
Compteur à lait	Production laitière	Pilotage du troupeau (santé, alimentation, performance)	-	500 à 1500€ par poste ou intégré au robot
Thermomètre	Température ruminale ou auriculaire	Troubles infectieux	-	5 à 10k€
	Température vaginale	Vêlages	Se = 100%, VPP = 100% (pour l'expulsion de la poche des eaux) ⁽²⁾	3 à 4 k€ (+ abonnement éventuel)
Bolus ruminal	pH ruminal	Troubles métaboliques	-	490€/bolus + 2500€ station de réception
Paramètres biologiques sur les produits				
Analyseur de lait	Concentration en cellules somatiques Conductivité (+Couleur)	Mammites	Se = 40 à 100% , Sp = 18 à 99%	7500€ + consommables 500 à 1500€ par poste ou intégré au robot
	Matière grasse, matière protéique	Troubles métaboliques	-	-
	Dosage de progestérone, Lactate déshydrogénase, β-hydroxy butyrate, urée	Chaleurs Mammites Troubles métaboliques	Se = 93 à 99%, Sp = 99% ⁽³⁾ Se = 80 à 82%, Sp = 98% -	40 à 50 k€ (+50 €/vache/an)
Paramètres morphologiques				
Balance	Poids	Troubles métaboliques Gestion de l'alimentation, Performances (allaitant)	-	5 à 10 k€
	Poids, Gain Moyen Quotidien, Buvées	Troubles de santé Gestion de l'alimentation	-	-
Paramètres comportementaux				
Ceinture abdominale	Contractions abdominales	Vêlages	-	3,5 à 4 k€
Podomètre	Activité motrice (nb de pas)	Chaleurs	Se = 59 à 96% , Sp = 96 à 99%	4,5 à 8 k€(10 k€ si tous animaux équipés)
	Position debout/couché	Vêlage Boiteries, Chaleurs	-	4,5 à 8 k€(10 k€ si tous animaux équipés)
Accéléromètre	Activité Physique	Chaleurs Boiteries	Se = 50 à 99%, Sp = 90 à 100%	4,5 à 8 k€ (+ abonnement éventuel)
	Temps d'ingestion Temps de rumination	Troubles de santé	-	
	Mouvements de la queue	Vêlages	-	3,5 à 4 k€
Détecteur électronique de chevauchement	Acceptation du chevauchement	Chaleurs	Se = 88%	6,5 k€
Microphone	Temps de rumination	Troubles métaboliques	-	8 à 10 k€ (tous animaux équipés)
Tapis	Pression de chaque patte	Boiteries	Se = 51 à 100%, Sp = 57 à 99%	-

⁽¹⁾ Se : Sensibilité ; Sp : Spécificité ; VPP : Valeur Prédicative Positive. Cf 2.2 pour leurs définitions. D'après Rutten *et al.*, 2013.⁽²⁾ D'après Chanvallon *et al.*, 2012⁽³⁾ D'après Friggens *et al.*, 2008

Thèse de Doctorat

Pierre CLEMENT

Titre de thèse : Valeur informative de l'activité générale et de la durée de rumination pour la détection des troubles de santé et le phénotypage chez la vache laitière

Title of thesis: Informative value of general activity and rumination time for the detection of health disorders and phenotyping in dairy cows

Résumé

De nombreux outils de monitoring ont été développés ces dernières années pour assister les éleveurs bovins dans la détection d'événements physiologiques (vélages, chaleurs) ou pathologiques (mammites). L'objectif de cette thèse était de produire des connaissances permettant de documenter l'intérêt de l'activité générale et des durées de rumination mesurées par un capteur destiné au monitoring des vaches laitières (HR-Tag, SCR Engineers Ltd, Israel), ou d'alertes générées à partir de ces données.

Une première partie s'est intéressée à l'intérêt des durées de rumination pour prédire les quantités ingérées durant la lactation. L'ajout des durées de rumination à un modèle de prédiction des quantités ingérées (NRC 2001) améliorait significativement les prédictions, mais le gain réalisé ne permettait pas d'accéder à un phénotypage précis de l'ingestion des animaux. Dans une deuxième partie, la valeur informative d'alertes générées, à partir des données d'activité et de rumination, pour la détection des troubles de santé a été évaluée. Trois méthodes de référence complémentaires ont été utilisées successivement pour déterminer l'état de santé des vaches : la détection par observation visuelle de l'éleveur, le repérage de chutes majeures de production laitière, et un examen clinique vétérinaire des animaux. Les résultats indiquent que même si les alertes de détection des chaleurs générées par le système de monitoring sont fiables, les performances de détection des algorithmes développés pour la détection des troubles de santé sont insuffisantes pour que les éleveurs puissent se reposer entièrement sur un tel système d'alarme pour leur détection.

Mots clés

Bovin, monitoring de la santé, élevage de précision, épidémiologie, algorithmes, cusum, rumination, activité

Abstract

In recent years, many monitoring tools have been developed to assist farmers in detecting physiological (calvings, heats) or pathological (mastitis) events. The aim of this Ph.D. thesis was to generate knowledge on the interest of general activity and rumination time measured by a sensor designed for dairy cow monitoring (HR-Tag, SCR Engineers Ltd, Israel), or of alarms generated based on these data.

First of all, the interest of rumination time for dry matter intake prediction during lactation was investigated. The addition of rumination time to an existing dry matter intake prediction model (NRC 2001) significantly improved the prediction, but the gain was not sufficient to enable accurate intake phenotyping. Then, the informative value of alarms generated from activity and rumination data for the detection of health disorders was evaluated. Three complementary reference methods were used successively to determine the health status of the cows: detection of health disorders by visual appraisal by farm staff, identification of major drops in milk production and clinical veterinarian examination of the animals.

The results of this work show that even if heat alarms triggered by the monitoring device are reliable, the performance of the algorithms developed are not good enough for farmers to delegate entirely health disorders detection to the alarm system.

Key Words

Dairy cattle, health monitoring, precision dairy farming, epidemiology, algorithms, cusum, rumination, activity