Use of MIR spectral data of milk in the detection and prevention of lameness in dairy cows

Auteur : Mineur, Axelle
Promoteur(s) : Gengler, Nicolas; 3041
Faculté : Gembloux Agro-Bio Tech (GxABT)
Diplôme : Master en bioingénieur : sciences agronomiques, à finalité spécialisée
Année académique : 2016-2017
URI/URL : http://hdl.handle.net/2268.2/3096

LIÈGE université
Gembloux
Agro-Bio Tech

Use of MIR spectral data of milk in the detection and prevention of lameness in dairy cows.

Axelle Mineur

Travail de fin d'études présenté en vue de l'obtention du diplôme de
Master Bioingénieur en Sciences Agronomiques

Année académique 2016-2017

Co-promoteurs : Prof. Nicolas Gengler, Prof. Hans Sölkner
Maître de stage : Dr. Christa Egger-Danner

Firstly, I would like to express my sincere gratitude to Prof. Nicolas Gengler, Prof. Johann Sölkner (BOKU) and Dr.Christa Egger-Danner for their continued guidance and support, as well as for sharing their immense knowledge in their respective fields and their insightful comments. Their guidance has been crucial for realising this master thesis.

I also take this opportunity to thank Prof. Johann Sölkner and his team at the Institut für Nutztierwissenschaften from the University of Natural Resources and Life Sciences, Vienna (BOKU) for welcoming me heartily in their department over the last 5 months.

I am very grateful to Dr. Clément Grelet for sharing his knowledge on the analysis of MIR spectra, using Matlab and the PLS-Toolbox and for making himself available to answer my specific questions.

Thank you to Dr. Astrid Kock who took the time to share her expertise in SAS with me.

Many thanks to Mr Robert Pesenhofer for sharing his expertise in hoof trimming in the field.

I am also grateful to the Centre Wallon de Recherches Agronomiques (CRA-W) for their training on 'Vibrational Spectroscopy and Chemometrics' which was so essential for this master thesis.

My sincere thanks goes to RINDERZUCHT AUSTRIA (ZAR) for providing the data I used in this master thesis.

Résumé

Actuellement, il n'existe pas d'équation de prédiction qui puisse détecter ou prévoir les problèmes de boiteries chez les vaches laitières, en se basant uniquement sur la composition du lait. Les boiteries représentent un sérieux problème de santé pour les vaches laitières, et ont des conséquences économiques pour les producteurs laitiers. L'objectif principal de ce travail consistait à tester dans quelle mesure il est possible de détecter problèmes de boiteries, à partir d'analyses de lait, à l'aide de la technologie moyen infrarouge (MIR). Celle-ci pourrait en effet procurer aux producteurs laitiers une méthode facile de détection précoce des boiteries. Les données sur les quelles ce travail est basé, ont été récoltées en Autriches et concernent les races Simmental, Brown Swiss et Holstein. Les indices de locomotion, allant de 1 (non-boiteux) à 5 (boiteries sévère), données aux vaches laitières, ont été organisées en une variable de classification non-boiteux - boiteux, avec un seuil fixé à 2.

Deux méthodes de calibration MIR ont été testées, visant à obtenir une équation de prédiction fiable pour la détection des boiteries. La première méthode repose sur une calibration MIR classique, et n'utilise que le spectre MIR comme variables de prédiction. Pour cette première méthode, l'obtention de meilleurs résultats a été atteint en opérant des sous-sélections de données de telle manière à réduire la variabilité des fichiers de données. La seconde méthode a utilisé, en plus du spectre laitier, des biomarqueurs dont les valeurs furent prédites à partir du spectre laitier. Les sensibilités et spécificités obtenues dépassaient rarement les 80 %, se situant la plupart du temps autour de 60 à 70 %. Ces résultats ne permettent pas une application sur le terrain. Néanmoins, ce travail suggère l'existence d'un lien complexe entre la composition du lait et les boiteries, au travers de blessures ou de maladies métaboliques et laisse entrevoir des perspectives pour des études futures.

Abstract

Currently, there is no prediction equation that enables the detection or prediction of lameness problems in dairy cows based on milk composition. Lameness is an important health issue for the animal and economic issue for the farmer. Therefore the general objective of this study was to test the feasibility of detecting lameness problems using mid-infrared (MIR) spectra from milk, as this could provide an easy and early detection method for the farmer. The data originated in Austria and was therefore distributed across three breeds; Austrian Simmental, Brown Swiss and Holstein. Locomotion scores, levels 1 (sound) to 5 (severely lame), given to the animals were organized into a classification variable non-lame or lame with a threshold of 2. Two different MIR calibration methods were tried with the aim of obtaining a reliable prediction equation for lameness. The first method was classic MIR calibration and used only the MIR spectra as predictive variables. In this case, trying to obtain better results was achieved through selection of subsets in such a way that overall variability would be reduced. The second method used extra predictive variables in the form of milk biomarkers for lameness to aim for a better prediction. The resulting sensitivities and specificities for both methods very rarely went above 80% and mostly seemed to stagnate around 60 to 70%. These results are certainly not high enough for application in the field, but this study does suggest the existence of a real, if complex, link between milk composition and lameness, through foot and claw lesions and metabolic disorders, which leaves the door open for further studies.

CHAPTER I: INTRODUCTION

Improvement of animal health and welfare goes hand in hand with the development of a more sustainable livestock agriculture. Healthier animals produce more and cost less, therefore farmers have always been concerned with the wellbeing of their animals. Bad health also leads to intensive use of antibiotics which enhances the risk of resistance appearing in bacteria. Moreover, the consumers are becoming increasingly concerned by the way their food is produced. This may lead to extreme actions taken by some animal welfare organizations investigating under cover to find shocking images. This provokes, not always deserved, criticism of farmers and other actors along the production chain. Farmers and especially dairy farmers sense the social scrutiny they are exposed to.

In this context lameness is a very important disease and this from different angles. It causes acute pain, distress and discomfort to the animals and can be an appalling sight for the general public. The related infections often tend to require an antibiotic treatment. Furthermore, it is the third costliest health related problem on dairy farms, after mastitis, metabolic diseases and their associated fertility problems. Reducing the prevalence of lameness is thus crucial for the wellbeing of the animals, for the profitability of the dairy farms, as well as their public image, and for public health reasons.

Reducing the prevalence of lameness requires improvements in the management of the farms and in the breeding of dairy cows. Before any actions can be taken, it is essential to be able to assess a lameness incident, at the early stages of its development, and to establish the lameness status of a given cow. Waiting until the moment the cow has become severely lame is not only ethically questionable, it also puts in jeopardy any chances of her recovery.

There are clear indicators of the fact that lameness is associated to physiological (e.g. from negative energy balance to inflammation) or behavioural (e.g. feeding habits) changes that should have repercussions on (fine) milk composition. Even if these facts have been recognized for a long time, assessing such milk based biomarkers remains very difficult and expensive.

An important recent development is the extended use of mid-infrared (MIR) spectrometry of milk. This technology has the potential to become an interesting ally in the detection of lameness and subsequently in the reduction of lameness levels on farms. The use of MIR is well established for major milk components. It is non-invasive and does not require any supplemental manipulation of the animal. Milk samples are routinely taken during milk recording, therefore using these samples will increase neither the stress levels of the animals nor the workload of the farmer. As a matter of fact, MIR spectrometry based analysis of milk is already commonly performed on milk payment samples and for the cows in milk recording. Currently, the MIR analysis provides at least fat and protein contents. In many countries it is also used to obtain the levels of urea, lactose and even other novel traits such as fatty acids.

MIR spectrometry is based on the development of appropriate prediction equations. These equations are developed through a process called calibration. A calibration model links known reference values of the trait of interest to the MIR spectral data, which represent the absorbance of infrared light by the corresponding milk sample. Recently, animal scientists started to think further, linking MIR spectra directly to the "status" of the animal. The idea that lameness changes have repercussions on the milk composition supports the initial hypothesis that these changes can be detected through the use of MIR spectra. This directly links MIR spectra to the condition of being lame. This could be especially useful for early lameness detection, i.e. animals are already affected before showing clear clinical signs and are potentially more easily overlooked, or for large herds where it is harder to keep regularly track of all animals with the same level of detail. Any prevention of severe lameness needs the early detection of cows that might be in the process of developing a lameness problem. That would enable preventive treatment and avoid the problem altogether.

To our knowledge there were, until now, no studies linking the condition of being lame to MIR spectra. Therefore the general objective of this study was to test the feasibility of detecting lameness problems using MIR spectra from milk. This master thesis will address the topic in five chapters. After the introduction, the second chapter will give information about the general state of the art in matters of lameness, milk composition and MIR methodology. The third chapter will explain the specific materials and methods used in this research. The fourth chapter will report the results found and discuss them. A final conclusion and perspective chapter will put this thesis in a more general context and give directions for future research.

This thesis was written in partnership with the University of Life Sciences, Vienna (BOKU), were I resided and worked on it for 5 months, and with RINDERZUCHT AUSTRIA, who graciously provided the data for this work from their "Efficient Cow" project.

CHAPTER II: LITERATURE REVIEW

1. Introduction

This chapter, constructed around three main steps, summarizes the state of the art in matters of lameness, milk composition and mid-infrared spectroscopy of milk. The first step is to establish the condition of lameness, its evaluation, prevalence, causes and impacts. Then, the link between lameness and milk composition changes will be discussed. Finally, the concept of milk mid-infrared technology will be explained and its current uses discussed.2. Lameness

2.1. Definition

According to the Merriam-Webster dictionary (2017), lameness or the condition of being lame is defined as "having a body part and especially a limb so disabled as to impair freedom of movement" or being "marked by stiffness and soreness". This broad definition makes it clear that lameness is not a disease in itself, but a complex condition that can basically result from other problems such as injuries to the feet and legs; metabolic problems or similar issues. To get a better understanding of the multifactorial condition that is lameness, the rest of this chapter about lameness (2.) is composed of the assessment of lameness through locomotion scoring, its prevalence and in general and more particularly in Austria and the causes and impacts of cow lameness.

2.2. Locomotion scoring

Lameness is not only a condition that is potentially difficult to define, but also difficult to assess. When observing lame animals, it is essential to be able to score the severity of the condition with more refinement than only differencing immobility from freedom of movement. Correctly assessing lameness is crucial in every lameness reduction program. First, as explained by RUTHERFORD *et al.* (2009), farmers often tend to have an inaccurate perception of the level of lameness in their herds. Moreover, this can also become a serious problem when the aim is to reduce the economic impact of lameness on the farm's profitability. A commonly used strategy to assess lameness is to score locomotion (mobility) of the animals as a proxy of the gradual absence of lameness. Mobility is also interesting as it can be considered directly linked to the economic impact of lameness (e.g. production loss, veterinarian treatment costs). To this end, several locomotion scoring systems were developed. Most systems are very similar, for this reason we will further discuss more in detail only two of these, in particular the system developed by SPRECHER *et al.* (1997) and that of Manson and Leaver (CHANNON *et al.*, 2009).

All systems are based on the grouping of animals with similar locomotion behavior into classes following the example of the type linear scoring system that is widely used in livestock. Linear scoring systems are divided into classes, hereafter

often called severity or clinical levels, used. For example, SPRECHER *et al.* (1997) used a 5-point visual locomotion scoring scale, Manson and Leaver (CHANNON *et al.* 2009) added 4 intermediate levels: 1.5, 2.5, 3.5 and 4.5 leading to a 9 level scoring system. Most systems score in the direction of lameness, not mobility, so a higher locomotion score means increased lameness severity and worse mobility.

There is also some variation in what is scored by the two systems. Sprecher's system assesses the severity of lameness based on gait regularity and back posture of the cow during standing and walking, ranging from 1 (normal) to 5 (severely lame). Figure 1 shows pictures and associated descriptions of the 5 clinical levels of lameness according to this 5-point system.

	Lameness score	Clinical description	Assessment criteria
	1	Normal	The cow stands and walks with a level-back posture. Her gait is normal.
	2	Mildly lame	The cow stands with a level-back posture but develops an arched-back posture while walking. Her gait remains normal.
	3	Moderately lame	An arched-back posture is evident both while standing and walking. Her gait is affected and is best described as short-striding with one or more limbs.
	4	Lame	An arched-back posture is always evident and gait is best described as one deliberate step at a time. The cow favors one or more limbs/feet.
	5	Severely lame	The cow additionally demonstrates an inability or extreme reluctance to bear weight on one or more of her limbs/feet.

Figure 1: Locomotion Scoring System, table adapted from SPRECHER *et al.* (1997)

An important element in each scoring system is the gait. Different aspects of gait can be used to determine whether the observed cow has a normal or an irregular gait. Some can be more easily observed than others. In Sprecher's system the following elements are used. First of all, there is the reluctance to bear weight on one or more of the legs as opposed to a healthy cow that will distribute her weight evenly between her four limbs. This will often be combined with an obvious head bob as the cow uses the movement of her head to change the way weight is distributed between her different steps. A healthy, walking cow usually keeps her head quite steady or moves it very freely and smoothly, but does not have a jerky or strong head movement. The head bob can also be more subtle when linked to asymmetric steps. Asymmetric steps can be seen and heard as instead of having a regular 1-2-3-4-beat walk, one of her strides will be longer or shorter than the others. When looking at the legs, good joint flexion is important to develop a nice, smooth walk for the cow. If one of the cow's joints hurts, she will be more limited in her mobility and ability to flex her joints and extend her legs. This will result in keeping her leg straighter during the swinging phase of the movement and probably make a shorter stride. This can often result in not tracking up properly as the cow will take smaller steps. Tracking up is the fact of having the tracks of the hind legs fall on top of or very near the imprints of the front hooves. Shortening the stride and therefore not tracking up correctly is also a clear sign of lameness and it will be more or less pronounced for different levels of lameness, from being only 1 hoof length behind to 4 or even 5. The last gait characteristic linked to the cow's legs is whether she is swinging her legs, usually her hind legs, in or out. A pain-free cow's hind legs usually go forward in a straight line, but a lame cow may bring her legs forward by swinging it towards or away from her body in a semi-circular or elliptic motion. Finally, the back arch of a cow can tell us a lot about pain or discomfort she might be feeling as a healthy and comfortable cow will keep her back flat, from the shoulders to the hip bones, when standing and walking. When feeling a little discomfort, the cow might still keep her back straight when standing and arch it only slightly when walking as if she were walking on eggs. This would result in a lameness score of 2. However, should she suffer from greater discomfort in one or more of her legs or hooves, she will start to show a more pronounced arch when walking. The scoring system by Manson and Leaver (CHANNON *et al.*, 2009) as most other systems uses very similar definitions. However there are differences in the system by Manson and Leaver as every cow is observed for a duration of 30 seconds, and this during different types of activities, e.g. turning.

There are many possible uses of locomotion scores (Zottl et al. 2017; Groen et al. 1997). The most relevant is the use of the developed fine phenotypes to get early warnings to farmers and veterinarians on deteriorating mobility of a given cow and to use these mobility scores in genetic evaluations. In many situations, successive lameness scores will need to be comparable for a given cow and across cows. To make sure the locomotion scores obtained are not biased, a few rules should be respected when locomotion scoring cattle. First of all, as the locomotion scoring systems look at back posture among other things, it is important to not score cows on slopes or slippery surfaces. The best surface to test lameness is a flat, level surface with adequate traction (Berry et al. 2017). Secondly, cows should not be running or pushed forward artificially by humans, but should be given the opportunity to show their natural walking gait. This is because accelerating might increase the difficulty to spot or hide altogether a mild lameness. Considering all of this, a nice timing for

scoring the herd might be just after milking, when cows leave the milking parlour spontaneously and at their own pace, provided that the corridor they move through is straight and long enough to have a clean rectilinear walk and broad enough to accommodate two cows so that faster cows can overtake slower ones instead of pushing them into a different pace.

2.3. Prevalence

2.3.1. In general

A few of these studies are listed here, in an order of increasing prevalence. In a UK study by Rutherford et al. (2009), around 1 in 6 to 1 in 5 cows (16.2%, 16.3% and 19.3% in the autumn, winter and spring observation) were found to be lame. Huxley et al. (2004) and Haskell et al. (2006) found quite similar results in their respective studies. The first studied organic herds and found about 24% of animals suffering from lameness. The second found 15% of lameness prevalence for grazing herds, but found more than double that for non-grazing herds at 39% lameness prevalence. A similar result was found by Barker et al. (2010) with 36.8% (range of 0-79.2%) of English and Welsh cows not being sound. However, an older study from 1997 by Sprecher et al. found a much higher lameness prevalence with 65.2% of cows being lame.

2.3.2. In Austrian dairy farms

The previously named studies were conducted on herds composed mostly of Holstein cattle. For this work however, the Austrian data was recorded on three important cow breeds: the Holstein, the Simmental or Austrian Simmental and the Brown Swiss or Brown Swiss. Austrian farms are also structured differently than Belgian or British ones with, on average, smaller numbers of cows per farm. However, claw horn lesions and associated lameness also have a big impact on farm profitability as claw or limb problems was the sixth most important reason for culling cows in Austrian dairy farms in 2011 (www.zuchtdata.at). It is therefore also interesting to look at the results of studies that took place specifically in Austria. ROUHA-MÜLLEDER *et al.* (2009) conducted a study in 80 organic and non-organic Austrian dairy farms that milk Simmental cattle/Simmental. Like BARKER *et al.* (2010), they also found that lameness prevalence varied greatly across the farms. Some farms did not possess any lame cows, while up to 77% of cows were affected in other farms. On average, they found a prevalence of 36.0%.

2.4. Causes of lameness

After describing in Chapter 2.2. how the lameness of a given cow can be assessed, this chapter will discuss the causes of lameness. Lameness is a complex and multi-factorial disease with mechanical, infectious or metabolic causes that are all interconnected. This chapter will shed some light on these different causes, and will be structured as follows: it

starts with a short description of the feet and legs of a cow, then it moves on to foot and claw disorders, and to metabolic disorders that can cause lameness.

2.4.1. Anatomy of a cow limb

In order to understand the origin of lameness, it is essential to understand the structure of a cow leg and foot. The hoof is composed of 2 claws, a medial or inside claw, and a lateral or outside claw. Two smaller, non weight bearing pieces of horn, called dew claws, are found at the height of the fetlock. Figure 2 shows a section of a cow claw.

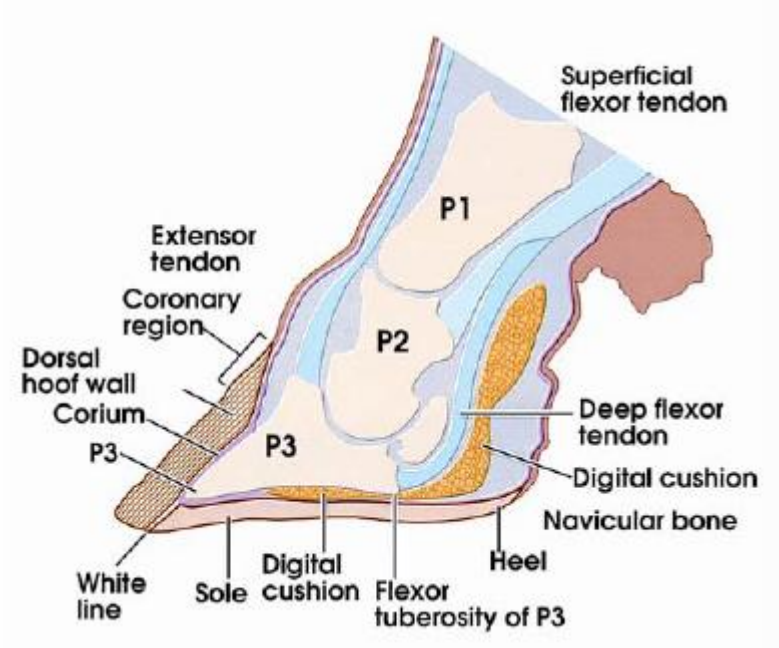

Figure 2: Anatomy of a cow claw (Gooch 2003)

P1, 2 and 3 are the pedal bones. P3 rests on the digital cushion, and is surrounded by a thin layer of corium (RODRIGUEZ & DEFRAIN, 2017). The digital cushion is an elastic shock absorber, composed of connective tissue, and varying amounts of fat. Having a healthy digital cushion is very important as this structure absorbs a lot of pressure at each step the cow takes (RÄBER *et al.*, 2004). The corium is rich in blood vessels and nerves, and it links P3 to the hoof wall, made of horn, through the corium laminae. Corium laminae look like baleen, the filter-feeder system found in the mouth of certain whale species. They vertically interweave with hoof laminae made of horn. This enables to firmly link the skeletal structure of the claw to the outer hoof wall, while still allowing the hoof wall to grow, and slide downwards. To the underside of the claw, the corium also links P3 to the sole (Anon 2017b). The white line, the junction of the sole and outer hoof wall, is a somewhat weaker spot, susceptible to penetration by stones and dirt (RODRIGUEZ & DEFRAIN, 2017). The fat contained in the digital cushion

and the corium laminae are two structures that will be important when metabolic sources of lameness are discussed in 2.4.2.

The weight distribution between the 8 claws of a cow is uneven. First of all, the front legs carry, on average, 50 to 60% of the weight of the cow, while the hind legs carry 40 to 50% (ANON, 2017a). Then there are the differences between medial and lateral claws. The front claws have a relatively equal weight distribution, tending slightly to the medial claw. For the hind legs however, this difference is much bigger. VAN DER TOL *et al.* were able to establish that a cow with trimmed feet carries about 70% on her lateral claw, and only 30% on the medial one. This difference can go up to 80-20% in untrimmed feet (2002, 2004). This big disparity puts more pressure on the outside claw, and may put it at higher risk of injury. This is in accordance with studies by CLARKSON *et al.* (1996) who found almost four times more lesions on the outer hind claw, than on the inner one.

According to CLARKSON *et al.* (1996), HERNANDEZ *et al.* (2002) and KOFLER (2014), 80 to more than 90% of lameness problems can be traced back to problems of the hooves, or the skin around it. However, a cow foot is not the only part of her limb that can suffer from injuries. The rest of the leg, especially the joints are sensitive to sprains and trauma, occasioned by either blunt force, or a penetrating wound (DALY 2014a). CHAPINAL *et al.* (2013) did a study across 53 dairy herds in the USA, and found an average prevalence of 58% (st d of 31%) for hock injuries, and of 16% (st d of 15%) for knee injuries. Other joints susceptible to problems, are the stifle and the fetlock (DALY, 2014a). These joints are shown on Figure 3.

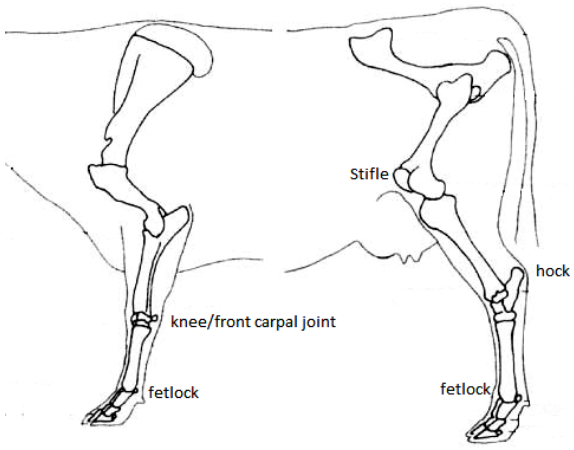

Figure 3: Joints of cow legs (ASHWOOD, 2011).

On some occasions, lameness can also find its cause in the lower back. For example, if a cow is mounted by another, heavier individual, it can cause nerve damage and general weakness that may translate into lameness (DALY, 2014a).

In the next two parts of this chapter, both wounds and metabolic problems will be discussed as causes of lameness. For the purpose of a clear explanation, they were separated into two distinct categories. However, it is important to note that foot and claw disorders are interconnected with metabolic disorders: a cow having one of the two problems is also at higher risk of developing the other one (ESPEJO *et al.*, 2006; VERMUNT, 1992). This is explained in chapter 2.5 which deals with the impacts of lameness.

2.4.2. Foot and claw disorders

2.4.2.1. Classification, description and prevalence of lesions

Foot and claw disorders are very complex, and not easy to classify. The ICAR claw health atlas (EGGER-DANNER *et al.*, 2015) describes, and depicts the most important disorders. As a matter of fact, there exist slightly different ways of classifying foot and claw disorders, also referred to as lesions, according to different authors. For example, KOFLER (2015) makes a distinction between claw lesions, e.g. white line disease, and lesions of the skin around it, e.g. phlegmon, while DALY (2014c) or GARRY (2017) do not recognize this distinction. What all three of them do agree on, is the distinction between infectious and non infectious lesions, or injuries. Infections are caused by bacteria, while injuries have a mechanical origin. Injuries can sometimes lead to infections, as a tear or open wound in the hoof structure open the door to dirt and bacteria. Table 1 describes the lesions most often encountered in practice, and in literature.

Table 1: Some of the most common foot and claw disorders.
Descriptions are from the ICAR Claw Health Atlas (EGGER-DANNER *et al.*, 2015)

Lesion type	Lesion name	Lesion description
Claw injury	- Sole ulcer	Penetration through the sole horn exposing fresh or necrotic corium because of inflammation
	- Sole haemorrhage	Diffused and/or circumscribed red or yellow discoloration of the sole and/or white line
	- White line disorder	Gap between the sole and the wall often filled with faeces or decayed horn masses
	- Horn fissure	Horizontal crack in the claw wall
	- Double sole	Two or more layers of under-run sole horn
	- Corkscrew claws	Any torsion of either the outer or inner claw. The dorsal edge of the wall deviates from a straight line
	- Sole penetration	Penetration through the sole horn because of a foreign object
Claw infection	- Phlegmon	Symmetric painful swelling of the foot commonly accompanied with odorous smell with sudden onset of lameness
	- Dermatitis digitalis	Infection of the digital and/or interdigital skin with erosion, mostly painful ulcerations and/or chronic hyperkeratosis/proliferation
Skin injury	- Heel horn erosion	Dissolution and decay of the horn on the bulbs of the heel
	- Limax	Interdigital growth of fibrous tissue

Just like the prevalence of lameness varies a lot across studies, so do the prevalence of foot and claw lesions. This is due to the many environmental and other factors, discussed in 2.4.4., that may increase the likelihood of one lesion over another in different settings. However, when looking at the table in Appendix 1, some patterns start to emerge. This table combines the results of six distinct studies, and when considering types of lesions with a

prevalence above 10%, it seems that sole ulcer, dermatitis digitalis, phlegmon, and white line disease are quite prevalent across more than one of these studies.

2.4.2.2 Importance of pain in the development of lameness

It is important to make a clear distinction between the terms lesion and lameness; a cow may very well have a lesion at one of her hooves, and still walk soundly. A hoof lesion will provoke lameness if it provokes pain, and the cow will try to reduce the pressure put on the lesion, by shifting her weight. This notion was studied by WHAY *et al.* (1997; 1998). In 1997, they studied 15 Holstein heifers as from 2 months before parturition, until 2 months after parturition. In their study, lesions were scored by using 3 characteristics: the size, the severity, and the position, i.e. the region of the sole where they appeared. They found that the presence of sole lesions around the time of parturition, was common to all heifers, but only 7 of them developed an associated lameness. The main factor that determined if a lesion caused lameness was not its size, but its severity. A mild haemorrhage was unlikely to provoke lameness in the cows they studied. In a follow-up study, the same authors WHAY *et al.* (1997) compared the nociceptive thresholds of 42 sound dairy cows, with 53 lame dairy cows. Hyperalgesia is a state of increased sensitivity to pain. This sensitivity can be measured using a nociceptive threshold test. WHAY *et al.* (1998) found that lame cows were indeed in a Hyperalgesic state, i.e. possessed lower nociceptive thresholds, than sound cows. This shows that lameness is closely linked to the pain a cow might experience from a lesion, or another mechanical problem. However, this also implies that a cow may not show any signs of lameness, because of a small lesion or at the onset of a problem, if she does not experience any pain from it. Therefore pain is the cause of lameness, not a lesion.

2.4.3. Conditions linked to metabolic disorders

Metabolism is the combination of all the complex and incessant processes of transformation of matter, and energy of the cell or organism, during anabolism (organic edification), and catabolism (organic degradation) (LAROUSSE 2017). Therefore, a metabolic disorder results from a disturbance in the balance of normal chemical reactions of the metabolism. Metabolic disorders have an ambivalent cause-consequence relationship with many conditions, as well as with lameness. In this part of the literature review, the focus will be on metabolic disorders as a cause for lameness.

Literature converges to three main conditions linked to lameness, that will be discussed in particular. Two of these are associated with metabolic disorders, while the third is a metabolic disorder. The first condition is loss of Body Condition Score (BCS) due to (excessive) body fat mobilisation, and its repercussions on hoof structure. It is strongly linked to the metabolic disorder, ketosis. The second condition is laminitis. Laminitis is not a metabolic disorder in itself, but is linked to the metabolic disorder called subacute ruminal acidosis or SARA. Researchers are not yet sure about the exact causes of laminitis, as it is a multifactorial condition, but SARA is often named as a possible cause (COOK *et al.* 2004; OETZEL 2015; GOOCH 2003; GARRY 2017). In this work, we will focus on laminitis caused by

SARA. The last and smaller part will explain why milk fever, a metabolic disorder, is a possible cause of lameness.

2.4.3.1. Body fat mobilisation

The Body Condition Score (BCS) is a visual estimation of the body fat reserves of a cow, and is therefore an excellent tool to help monitor the condition of the cows (BRAUN *et al.* 1986), and is more or less a common tool in herd management (ZOTTL *et al.* 2017). According to METZNER *et al.* (1993), on a BCS scale from 1 to 5, a dairy cow should ideally have a score of 3.50 (3.25 - 3.75) at the time of calving. This score might then drop slightly during the first part of the lactation, but by the time of drying, at the end of the lactation, the cow should have regained her weight, and refilled her bodily reserves. The cow should again have a score around 3.50. The aim is to have a good condition, neither too fat, nor too weak.

If the cows lose too much body condition, they risk ending up with a negative energy balance. When that happens, their bodies will try to compensate by burning protein and fat reserves, and they may start suffering from ketosis (COLLARD *et al.* 2000). Ketosis is a metabolic state during which some of the body energy supply comes from ketone bodies (HERDT 2017). Body fat mobilization uses fat reserves from adipose tissues all around the body, and transforms them back into fatty acids which are transported to the liver, and transformed into ketone bodies, hence the name of this condition.

In many studies, the lowest BCS has the highest risk of developing lameness. This is because ketosis also mobilizes fat reserves of the digital cushion, largely constituted of fat. This induces a thinning of the digital cushion, and a reduction in its efficiency as a shock absorber, which facilitates damage to the corium (ESPEJO *et al.*, 2006; MULLING AND GREENOUGH, 2006, cited by DIPPEL *et al.*, 2009). Indeed, BICALHO *et al.* (2009) and NEWSOME *et al.* (2017) found that BCS was positively associated with Digital Cushion Thickness (DCT).

The study by BICALHO *et al.* (2009) also found DCT to be a strong predictor of lameness, as cows in the upper quartile of DCT showed a 15% lower chance of lameness than the lower quartile. This illustrates the real importance of a well formed and supportive digital cushion to provide an optimal dissipation of forces within the bovine claw, to prevent lesions (RÄBER *et al.*, 2004).

Indeed, studies by BICALHO *et al.* (2009) and NEWSOME *et al.* (2017) found out that sole ulcers, white line diseases and sole haemorrhages were significantly associated with thickness of the digital cushion, and that they had a thinner digital cushion before the lesion occurrence. Interestingly, NEWSOME *et al.* (2017) observed a thickening of the digital cushion after the lesion occurred, possibly showing an increased vascularisation, oedema, or inflammation.

As explained previously, the association between a lower BCS and more lameness, is quite straightforward. Nonetheless, some authors point out that being too fat would not be an advantage for the cow either. Indeed ROUHA-MÜLLEDER *et al.* (2009) found an association of higher body weight and more cases of clinical lameness. There are two reasons for this. On the one hand, being heavier puts more pressure on the claws, which is especially important when standing on hard surfaces for long periods of time. On the other hand, a higher percentage of fat cows usually means higher milk yields, and high milk yield is a risk factor for lameness (ROUHA-MÜLLEDER *et al.*, 2009; ARCHER *et al.*, 2011). DIPPEL *et al.* (2009) also mention it, although it was not the case in their study as they only had very few cows above a BCS of 3.75.

2.4.3.2. Laminitis

Laminitis is an inflammation of the sensitive laminae of the corium (GOOCH, 2003). The aetiology of the disease is not yet precisely known, and many factors can influence the onset of laminitis (VERMUNT, 1992). Two main types of causes are generally accepted: nutritional ones, and environmental ones (COOK *et al.*, 2004; GOOCH, 2003). The nutritional cause that is most often mentioned in literature, is Sub Acute Ruminal Acidosis (SARA) and it will be the one discussed here (COOK et al.,2004).

SARA happens when the rumen content experiences a drop in pH below approximately 5.8% (GOOCH, 2003). This can happen as a consequence of feeding a diet too rich in carbohydrates, but too poor in fibres (GOZHO, 2005). The epithelial cells of the rumen are not protected by mucus, and the drop in pH makes them suffer a chemical damage, called rumenitis. This causes an inflammation of the epithelium, and bacteria colonize the epithelial cells and even leak into blood circulation. It seems that some of these bacteria and other released substances, trigger changes in the corium's blood vessels which causes inflammation and haemorrhages (OETZEL, 2015).

In the case of acute or subacute laminitis, the effects of this inflammation are very rapid and severe, and cause a separation of the hoof laminae and the corium laminae. This causes instability of the third pedal bone, which is free to slip downwards, and puts extra pressure on the sole. This is very painful for the cow, and causes lameness (MASON, 2008).

However, many cows experience subclinical laminitis, which results in the growing of softer hoof horn, and leads to a heightened risk of secondary claw lesions from concussion, like sole ulcers or white line diseases (VERMUNT, 1992). Usually, the effects of hoof lesions only become visible 6 to 8 weeks after the SARA episode. (GARRY, 2017; GOOCH, 2003).

Environmental factors linked to housing and management, that may influence the onset of laminitis, are excessive standing on concrete which can be related to uncomfortable stalls, over-crowding, heat stress, and long parlour-hold times (more than 3 hours per day) (GOOCH, 2003).

2.4.3.2. Milk fever

Milk fever is a possible clinical manifestation of hypocalcaemia, but not all cows with hypocalcaemia develop milk fever (GOFF, 2008). Hypocalcaemia is defined by the Merriam-Webster Dictionary as a deficiency of Calcium in the blood. Milk fever has a strong effect on the health and behaviour of the animal, namely reduced appetite, cold body skin, weakness, lack of coordination, and being unable to stand, are all symptoms of this ailment (SOGSTAD *et al.* 2006).

SOGSTAD *et al.* (2006) found that milk fever was positively associated with moderate and severe haemorrhages, and all levels of sole ulcers. A possible explanation is that a cow suffering from milk fever, and struggling to stand up, puts uneven or increased pressure on her claws. This mechanical pressure can cause lesions or worsen subclinical laminitis. This study also cited MÜLLING *et al.* (1999), who reported that plasma calcium concentrations have an influence on differentiating epidermal cells, including those producing horn, which could cause the formation of dyskeratotic horn. Dyskeratotic horn is more fragile for it is constituted of abnormal, premature, or imperfect keratinocytes, which are the dead cells that constitute the horn. Weaker hoof horn is more susceptible to damage, and therefore makes the animal more susceptible to lameness.

2.5. Impact of lameness

If left untreated, lameness problems can have a considerable impact, not only on the productivity of a cow, but also on her health, behaviour and welfare.

2.5.1 On cow health, fertility, behaviour and welfare

As was already mentioned, there exists an ambivalent cause-consequence relationship between metabolic disorders, and lameness. Here we will discuss metabolic disorders as a result of lameness. Indeed, a cow that does not have a metabolic problem, can hurt herself or get infected. As a consequence, the cow develops a foot lesion with lameness of a mechanical origin.

Foot and claw disorders are among the most painful ailments cows can suffer from, and which in turn influences their behaviour (WHAY *et al.*, 1997). This discomfort decreases the willingness of the cow to move, and as a result, she will lie down for longer periods of time (ITO *et al.* , 2010; KING *et al.*, 2017). Severe lameness can even cause the cow to lie down so long, that it reduces the feeding time and dry matter intake of the cow (DIPPEL *et al.* 2009). Therefore, she takes in less energy, which can be a serious problem for the highly productive dairy breeds we have today, as it can result in loss of body condition (ESPEJO *et al.*, 2006).

This is especially true for cows who just calved, and are at the start of their lactation. At that period in time, the calving and the onset and increase of lactation, cost a lot of energy, and the cows find themselves in a negative energy balance, and lose body weight, like shown on Figure 4. In this state, a reduction of energy intake because of an unwillingness to move can have dramatic impact, and lead to metabolic problems like ketosis (COLLARD *et al.*, 2000).

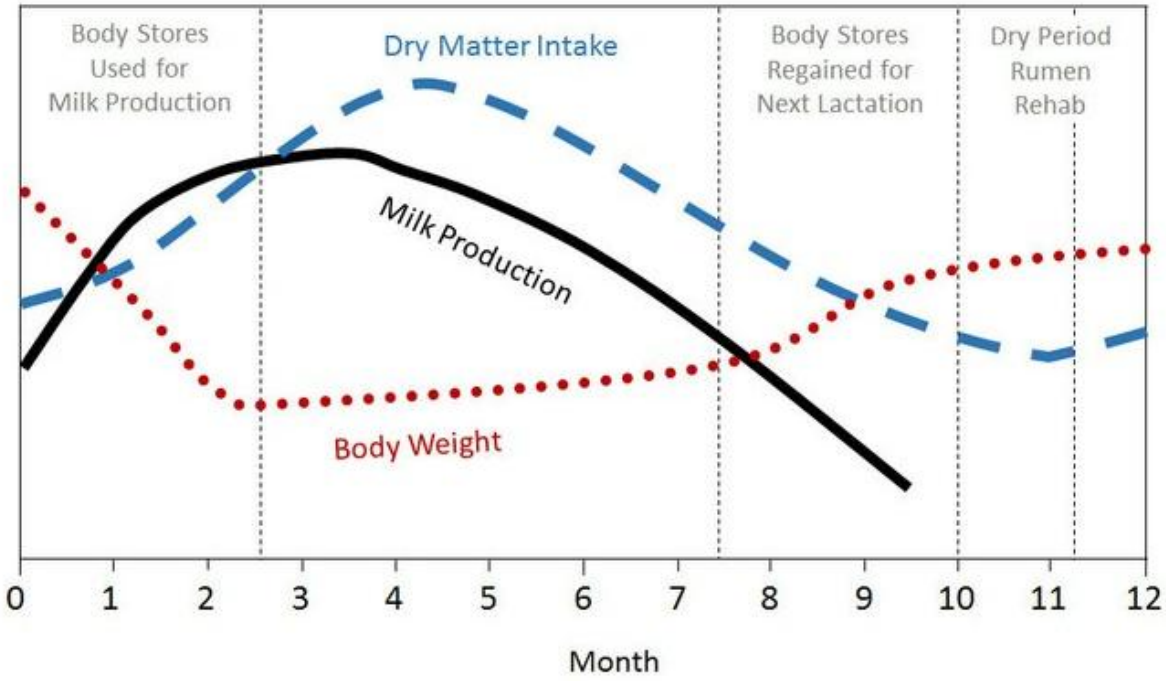

Figure 4: Changes in body weight, dry matter intake, and milk production over a single lactation (Braun et al. 1986).

KOFLER (2013) reported the incidence of lameness according to lactation month (Figure 5). As expected, metabolic changes at the moment of calving and at the start of lactation, have as consequence that most lameness cases appear during the first months of lactation.

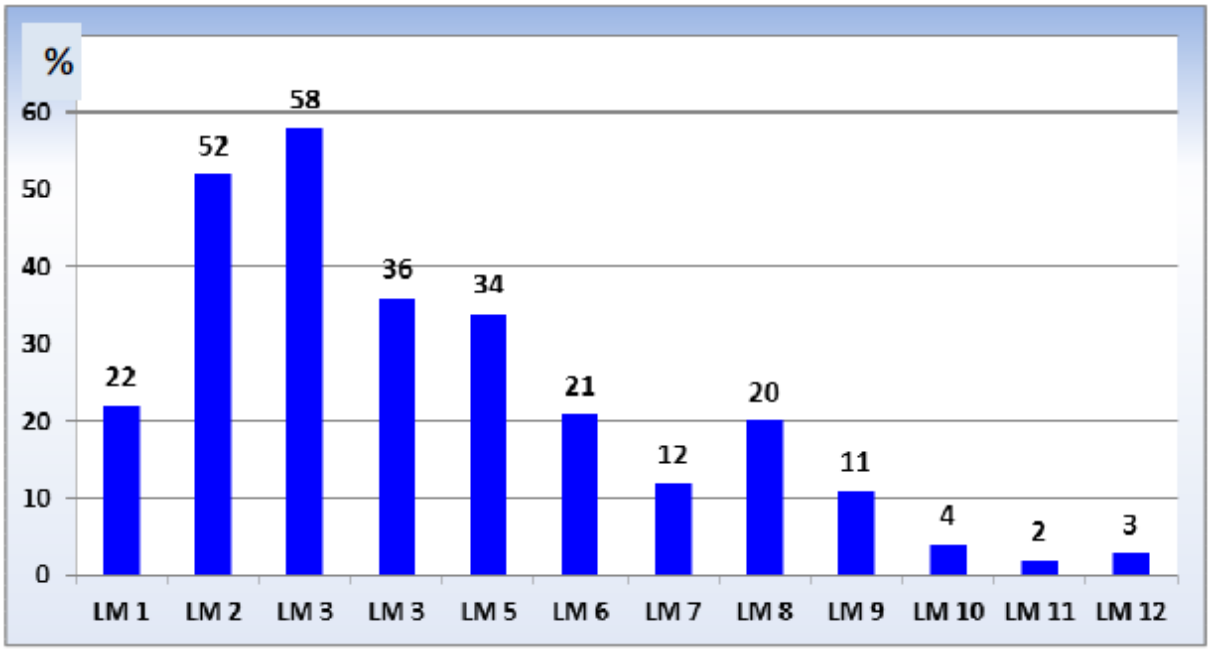

Figure 5: Incidence of lameness according to lactation month (LM) (KOFLER, 2013).

SOGSTAD *et al.* (2006) also linked lameness to a higher risk for developing mastitis. A possible explanation for this association, might be teat injuries due to the struggle to stand up from a lying position. These teat injuries can then lead to mastitis. However, this association between lameness and mastitis might not be strictly cause-consequence. Another explanation is that both diseases are related to cow comfort, management and hygiene, and therefore have a lower or higher incidence in similar management systems.

Fertility is usually among the first systems that shuts down, when an animal has health problems. This is also the case with lameness (SPRECHER *et al.*, 1997). Lesions such as heel-horn erosions, sole haemorrhages and ulcers, and white-line fissures were associated with longer calving interval in pluriparous cows, and longer calving and calving to last service intervals in heifers (SOGSTAD *et al.*, 2006; CHARFEDDINE & PÉREZ-CABAL, 2017). This again shows the importance of early detection and treatment to improve reproduction.

Another effect of lameness is its impact on the interaction of the cow with other individuals. When she is weaker, she sinks in the hierarchy, and can get bullied more easily by other individuals (KOFLER, 2013). This can be a stressful situation for the cow. By trying to escape bullies, she might slip and end up with additional contusions. Stress is also a triggering factor for the development of SARA.

Because of all these negative effects lameness has on cow health, the pain caused by it and the fact that these problems can last for weeks or even months (KOFLER, 2013), lameness is a major welfare problem. Even in the treatment of lameness lies a welfare problem. Indeed, pain control is virtually unknown. Treatment of lesions happens with surface antibiotics after cleaning the wound, which can lead to bleeding and be excruciating for the cow. Analgesics could be used to relief the pain, but are virtually never used. This poses an ethical problem as, even though the animal is clearly suffering, it does not receive pain control medication. It also shows a clear distinction with pets, who always receive analgesics in case of severe pain. However, with the increasing awareness of the public, this could also prove to be an important evolution in the treatment of lameness (ROLLIN, 2005).

So, despite being widely known as an important welfare problem, still, the prevalence of subclinical and clinical mastitis remain high. Farmers are mostly aware of the direct costs of lameness, because of the non-usable milk due to the waiting time for antibiotics. Lameness represents an increased work load for the farmer, and additional costs for hoof trimming, veterinary care and treatment. However, indirect costs do exist. These include: a reduction of the milk yield, reduction in fertility and associated culling, and longer calving intervals (CHARFEDDINE & PÉREZ-CABAL, 2017). Next to the welfare problem, a greater farmer awareness of the cost of lameness, could be a good incentive for them to aim at reducing this problem (BRUIJNIS *et al.*, 2010). The cost of lameness is therefore the next topic of discussion.

2.5.2. On profitability

The profitability of a dairy cow depends on her yield and on her longevity. The longevity in this case is not strictly speaking the duration of her life, but her capacity to stay in a good production rhythm, for example calving and start of lactation, insemination after three months, drying off after 10 months and next calving and start of lactation after 12 months. This pattern can be disrupted by disease, which can reduce milk yield and prevent the produced milk from being used because of antibiotic use, or an incapacity to become pregnant after first or following inseminations because of a loss of fertility. Lameness has these effects on the cow and is therefore an important problem (Sprecher et al. 1997; Charfeddine & Pérez-Cabal 2017; Bruijnis et al. 2010).

In fact, lameness is the third most important reason for economic losses in dairy farms, after mastitis in the first place and metabolic diseases and poor fertility in the second place (KOFLER, 2013; KOSSAIBATI & ESSLEMONT, 1997; BRUIJNIS *et al.*, 2010).

Economic losses can be divided into 2 kinds: losses of profit, and increases in costs. Losses of profit include: the reduction in milk yield, non-usable milk due to waiting time for antibiotics, culling and longer calving intervals. Next to this, increases in costs are due to an increased work load for the farmer, and additional costs for hoof trimming, veterinary care and treatment. The average cost of lameness, calculated in a recent Dutch study by BRUIJNIS

et al. (2010) gives a good idea of the economic loss that lameness can represent for a farmer. Table 2 shows the results they found.

Table 2: The average and minimum to maximum economic losses (in US-$/year), under the milk quota system of the Netherlands, for subclinical and clinical foot disorders, for a default farm with 65 lactating cows (BRUIJNIS et al., 2010).

Costs occasioned by:	Subclinical	Clinical	Total
Milk yield reduction	1219 (994 – 1.469)	949 (641 -1.349)	2168 (1.729 – 2.635)
Non-usable milk due to waiting periods		190 (39 - 377)	190 (39 - 377)
Culling		1084 (0 – 2.969)	1084 (0 – 2.969)
Longer calving interval	343 (207 - 477)	246 (145 - 373)	589 (415 - 767)
Extra work load for farmer		578 (353 - 832)	578 (353 - 832)
Extra hoof trimmer costs		148 (37 - 293)	148 (37 - 293)
Extra veterinarian costs		74 (0 - 240)	74 (0 - 240)
Treatment costs		67 (41 - 96)	67 (41 - 96)
Total	1562 (1.245 – 1.927)	3337 (1.736 – 5.435)	4899 (3.217 – 7.001)

The total average cost of $4899 is quite a substantial amount of money for a disease that can be controlled with good management. Per animal, this translates to an average cost of $75 per year. Milk production losses and culling were the most important reasons for economic loss.

BRUIJNIS *et al.* (2010) also looked at the costs of different foot lesions, and found that subclinical foot disorders cost an average of $18. This amount strongly increases for clinical disorders; $95 on average. This also encourages the idea that early detection and treatment, can have a strong effect on cost reduction.

The long-term effects of lameness can also be of consequence. WHAY *et al.* (1997) explained that an episode of clinical or subclinical laminitis, can have long-term effects on the animals. This is also the case for lesions as shown by CHARFEDDINE & PÉREZ-CABAL (2017) who did a study on data, provided by the I-SAP program, implemented by the Spanish Holstein Association to collect data on 6 claw disorders. Their data came from 804 Holstein dairy herds between 2012, and 2014 were collected by 25 trimmers. They found that an episode of severe sole ulcer or white line disease reduced the productive life of the cow by up to 71 days.

3. Milk composition and lameness

Animal breeding and management, and therefore any improvement through either management or selection requires relevant data that is as close as possible to the processes that are being assessed. This brings along the concept of biomarkers; "… objectively measured and evaluated … indicator of normal biological processes, pathogenic processes, or … responses to an … intervention" (NATIONAL INSTITUTES OF HEALTH). The definition of biomarkers may be very broad (e.g. STRIMBU & TAVEL, 2010). Therefore information collected at claw trimming can be considered as (early) biomarkers to the condition of lameness. Similarly, other visible changes in animal health and status, can be clinical (early) indicators of lameness such as a reduced feed intake, a deteriorating body condition (BCS) or weight loss.

But other invisible, subclinical, biomarkers or indicators for lameness may exist. (HAMANN & KRÖMKER, 1997) described that the status of a given cow can be assessed through the composition of three body fluids: blood, saliva and milk. Blood based biomarkers would of course be more direct in the case of lameness, but milk based biomarkers are more interesting, for they are much easier to assess on a large scale, and at reasonable costs (GENGLER *et al.*, 2016). This is where milk composition shows all its usefulness as a substrate containing many potential biomarkers. Furthermore, the biochemical milk profile is a reflection of the metabolic status of the cow (HAMANN & KRÖMKER, 1997).

In conclusion, there are clear indicators that lameness should be associated to physiological (e.g. from negative energy balance to inflammation), or behavioral (e.g. feeding habits) changes that should have repercussions on (fine) milk composition. In this chapter, we will describe several major and minor milk components that can be considered linked to lameness.

3.1. BHB, acetone, citrate

A first group of biomarkers for lameness is related to negative energy balance, and to hyperketonaemia. BHB is a ketone body. Acetone is a breakdown product of acetoacetate, another ketone body. Both are the results of lipomobilization when the cow falls into negative energy balance, which causes hyperketonemia (HAMANN & KRÖMKER, 1997). Hyperketonaemia is the rise in ketone bodies in the blood associated with subclinical and clinical ketosis.

Blood and milk acetone and acetoacetate in blood and milk were all highly correlated (phenotypic correlations), and acetoacetate in blood or milk gave the best sensitivity-specificity combination for the detection of subclinical ketosis in a study by ENJALBERT *et al.* (2001). BHB was also a good indicator if a little less then acetoacetate. In their study, a cow was considered subclinically ketotic if she had a blood BHB > 1200 µmol/l. Other authors

like McArt *et al.* (2012, 2013) also place this limit for hyperketonaemia at 1200 µmol/l blood BHB.

VAN DER DRIFT *et al.* (2012) found that genetic correlations between blood BHBA, often considered the reference and milk BHBA (0.52) or milk acetone (0.52) were moderate. However milk ketone bodies can be more routinely analyzed, so they could still be an alternative for breeding programs, trying to reduce hyperketonemia in early lactation.

Milk citrate was recently associated with negative energy balance, and recognized as a good early indicator for it (GRELET *et al.*, 2016; HAMANN & KRÖMKER, 1997).

Milk acetone was significantly correlated with energy supply for the first 4 months of lactation in a study by HAMANN & KRÖMKER (1997). In that study, acetone was in fact a better indicator of energy balance than the fat/protein ratio.

In conclusion, Milk BHB, acetone and citrate are good indicators of energy metabolism of single cows or whole herds, and potential indicators of clinical and subclinical ketosis (MCART *et al.*, 2012; HAMANN & KRÖMKER, 1997; VAN DER DRIFT *et al.*, 2012; ENJALBERT *et al.*, 2001; GRELET *et al.*, 2016).

3.2. Minerals

A second group of biomarkers is related to the levels of minerals in milk. Already HAMANN & KRÖMKER (1997) pointed out that the increase in milk concentration of some minerals like phosphorus and sodium, are indicators of calcium and phosphorus deficiency. TOMLINSON *et al.* (2004) explained that many molecules, including Calcium, Zn, Cu, Mn, vitamins A, D, and E, and biotin play an important role in the production and maintenance of healthy and strong hoof horn. Inferior quality of hoof horn increases the risk of developing claw disease. Moreover, potassium was recognized as a potential indicator for acidosis, this condition, and especially its subacute form (SARA) (1997), being linked again to lameness (COOK *et al.*, 2004).

3.3. Fat, protein

3.3.1. Fat

Lameness and fat content can either have a positive, or a negative association. If lameness is due to laminitis and therefore linked to SARA, then a reduction in milk fat can be observed, however this response is not always consistent in experimentally induced SARA, which can be explained by the fact that very short bouts of SARA might not have time to significantly influence milk content (PLAIZIER *et al.*, 2008).

If lameness is linked to ketosis however, then milk will see an increase in its fat content and fat percentage, because of a more important body fat mobilization (LOKER *et al.* 2012).

3.3.2. Protein

Several studies (e.g. LOKER *et al.*, 2012) showed that milk protein percentage was positively correlated with BCS. So, a lower BCS caused a lower protein percentage. This is because, even if negative energy balance triggers the mobilization of protein reserves, the mobilization of fat is considerably more important, which results in a decreasing milk protein percentage (LOKER *et al.*, 2012) .

High milk protein contents can be linked to a diet rich in protein, which can be a risk factor for laminitis because of SARA (DIPPEL *et al.*, 2009a). SARA causes an increase of acute phase proteins in blood, which is an indicator of inflammation, in this case, inflammation of the epithelial wall of the rumen (PLAIZIER *et al.*, 2008). Foot and joint wounds can also trigger an inflammation response with the release of acute phase proteins (HASKELL *et al.*, 2006).

3.3.3. Fat to protein ratio

Even if high fat to protein ratios can have other reasons, its increase at the beginning of the lactation is by some considered an indicator of an increased risk for negative energy balance. A high fat to protein ratio also shows significant negative correlation with BCS for the first 4 months of lactation (LOKER *et al,.* 2012). As explained earlier, both are indicators for increased lameness risks as in early lactation, the cow loses condition and mobilizes mainly adipose tissue. Energy provided by protein is comparatively much smaller. This translates into the fat and protein contents of milk, and explains the increase in fat to protein ratio in the milk (KOFLER *et al.*, 2013; HAMANN & KRÖMKER 1997; LOKER *et al.*, 2012).

3.3.4. Non esterefied fatty acids

Body fat mobilization during early lactation releases non esterefied fatty acids (NEFA). Part of these will be transformed into ketone bodies like BHB in the liver, but some of them remain as fatty acids. So, a certain concentration of BHB and NEFAs at the start of lactation is normal, but if this increase becomes too important, then this is an indication of excessive negative energy balance, and possibly ketosis during the transition period. If that happens, NEFA content in milk will also increase, making it an interesting indicator (McArt *et al.*, 2013).

When looking at specific NEFAs, VAN HAELST *et al.* (2008) saw that subclinical ketosis (SCK) provokes a decrease in medium-chain saturated FA, and an increase in long-chain FA in milk. The major NEFAs released are C16:0, C18:0, and C18:1 cis-9. C18:1 cis-9

is the most abundant and it's elevated proportions two weeks before subclinincal ketosis (SCK) diagnosis, makes it a good indicator of SCK.

4. MIR spectroscopy of milk

As shown in chapter 1.3. different components, often called biomarkers, in fine milk composition are very promising in their capacity to be used in the context of lameness. However acquisition of fine milk composition is expensive, therefore the need for large-scale and cheap phenotyping tools remains (DE MARCHI *et al.*, 2014). In this chapter we will describe infrared spectroscopy, and in particular mid-infrared (MIR), used in milk as a very useful technology with many advantages. The two major advantages are that, first, MIR is a fast method already used in routine for major milk components, therefore the logistics to take samples, to analyse them and to distribute results exist and secondly, MIR spectroscopy can provide a large amount of information since many molecules have a strong absorption (SMITH, 1996).

In the following paragraphs we will, first explain the general concept of spectroscopy, spectrometry, especially using infrared, then the use of MIR for the analysis of milk and the technologies and methods used to do this.

4.1. Spectroscopy

Spectroscopy is the study of the interaction between matter and electromagnetic radiation. This interaction, however, can vary in kind. The three most important types of interaction are absorption, emission and reflection (GENGLER *et al.*, 2016).

Smith (1996) explains how electromagnetic radiation is the appropriate term to describe light and in extension also infrared "light" as composed of an electric and a magnetic wave. These two waves undulate in planes which are perpendicular to each other, and go through repetitive motions called cycles. Therefore, a wavelength is the distance travelled by a wave during a cycle. The units of wavelength is distance per cycle (cm).

Not all electromagnetic radiation are visible or infrared, DE MARCHI *et al.* (2014) explains that the spectrum of electromagnetic radiation comprises different regions. Even if there is not an absolute consensus on their limits, these regions can be separated in function of their wavelengths: the x-ray region (~0.5 – ~10 nm), UV region (~10 – ~350 nm), visible region (~350 – ~800 nm), different infrared regions (~800 nm – ~1000 mm), microwave region (1 mm–1 cm), and radio frequency region (1 cm–1 m).

In spectroscopy however, the usual unit is not wavelengths, but wavenumber (W). Wavenumber is expressed in units of cycles per centimetre, abbreviated as cm^{-1}, because they are the inverse of wavelength as shown by equation 1 (Eq.1) (SMITH, 1996).

$$W = 1/\lambda$$

Wavenumber is therefore inversely proportional to wavelength. Expressed in wavenumbers, the infrared region of the electromagnetic spectrum is typically considered to range from

approximately 14 000 to 10 cm^{-1} and is split in the context of spectrometry into three distinct regions: the near infrared (NIR) from 14 000 to 4000 cm^{-1} (2.5 – 0.7 µm) the mid-infrared (MIR) from 4000 à 400 cm^{-1} (25 – 2.5 µm) and the far infrared from 400 to at least 4 cm^{-1} (1mm – 25 µm) (Smith 1996). As shown in Figure 6 different regions interact with different atomic or molecular events. If visible light and UV detect transitions of electrons, the different IR regions detect vibrations.

>14,000 cm^{-1} Visible & UV	14,000 to 4000 cm^{-1} Near IR	4000 to 400 cm^{-1} Mid-Infrared	400 to 4 cm^{-1} Far Infrared	< 4 cm^{-1} Microwaves
Electronic Transitions	Molecular Vibrations	Molecular Vibrations	Molecular Vibrations	Molecular Rotations

Higher Wavenumber	Lower Wavenumber
Higher Frequency	Lower Frequency
Higher Energy	Lower Energy
Shorter Wavelength	Longer Wavelength

Figure 6: Part of the electromagnetic spectrum (SMITH, 1996) used in spectrometry.

4.2. IR spectroscopy

As previously established, spectroscopy is the study of the interaction between matter and electromagnetic radiation. IR spectroscopy, also called vibrational spectroscopy because it allows detecting molecular vibrations, is mostly based on absorption spectroscopy but many other techniques might be used, especially reflectance when working with solids. IR spectroscopy based methods to quantify substances in solids (e.g. humidity in grain), liquids (e.g. fat concentration in milk) or gasses (e.g. methane in the air) are well established. Because of the quantitative aspect this type of spectroscopy is called IR spectrometry.

4.2.1. Methodology

The basics of IR spectrometry can be traditionally resumed (GENGLER *et al.*, 2016) as:

- Preparing a sample: this might imply grinding or homogenizing
- Guiding a beam of infrared light through the sample

- Recording it after passage / reflection: absorption or attenuation occurs when the frequency of the IR beam is equal to the vibrational frequency of a chemical bond or collection of bonds.
- Detecting through the comparison of emitted and recorded amount of energy absorbed / attenuated at each wavenumber (or wavelength) which? give the spectrum. These are recorded as transmittance. In traditional absorption spectrometry absorbance (A) is a reciprocal logarithmic function of transmittance: $A = -\log_{10}(T)$.
- This is repeated for each wavenumber across the wavelength range:
    - One wavenumber at a time using a monochromator (method used previously).
    - The entire wavelength range using a Fourier transform based instrument, the spectrum is generated after using the Fourier transform algorithm.
- The absorbance values are combined, mostly, linearly to generate a prediction of the trait of interest. These prediction (also called calibration) equation coefficients were computed *a priori* during a process called calibration.

4.2.1. Fourrier Transform Infrared (FTIR)

Many modern IR spectrometers are FTIR spectrometers which have the particularity of containing an interferometer and use Fourrier Transform based technologies. The interferometer is composed of a beamsplitter and two mirrors. A beamsplitter is a particular type of mirror that reflects one part of a light beam and lets the other through. So instead of having one light beam that travels to the sample and whose absorption by the matter is then measured, the light beam is split into two. These two light beams each travel to a mirror, one to an immobile mirror, the other one to a moving mirror. Both mirrors then reflect both beams and they are recombined into one beam at the beamsplitter before being sent out towards the sample. After the light has interacted with the sample, the detector registers an interferogram. This process is shown on Figure 7 that shows a Michelson interferometer which is the most common type. The Fourier transform, a mathematical operation from Fourier's theorem, is used to translate an interferogram into an infrared spectrum.

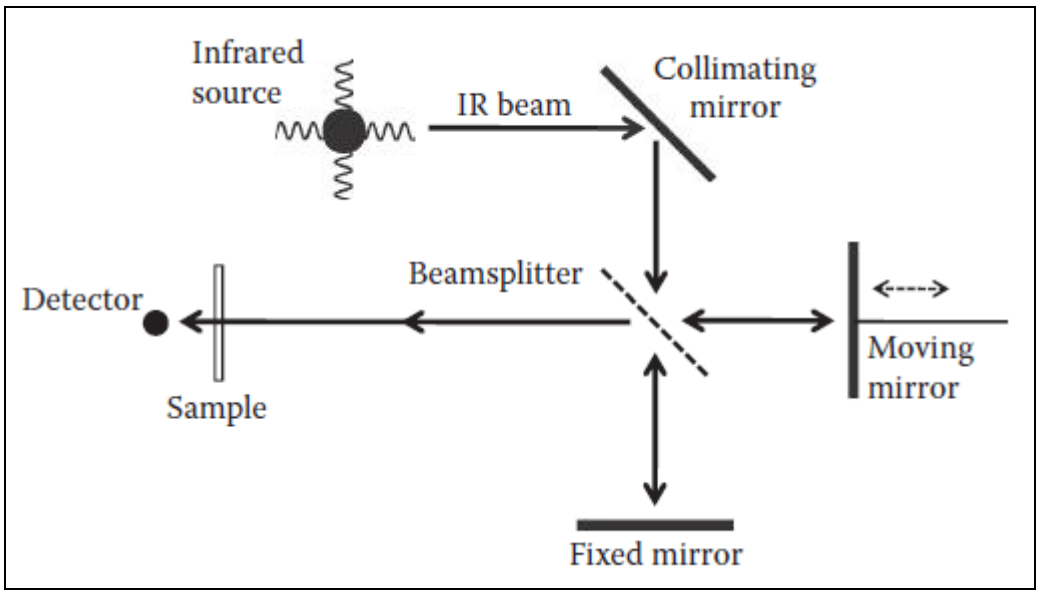

Figure 7: The optical diagram of a Michelson interferometer (SMITH 1996).

The main advantage of the FTIR is that the moving mirror enables the user of an FTIR to obtain the complete MIR spectrum for a sample, containing all wavenumbers, in one go. This is because when the two light beams recombine, they interfere and can form a large array of amplitudes and therefore wavenumbers. This means that their amplitudes combine into a much bigger amplitude than the two separated beams (constructive interference) or into a much smaller amplitude than the two separated beams (destructive interference). An example of both types of interference is given in appendix 1.

Another advantage of the FTIR is the better signal to noise ratio (SNR), which means a reduction in noise. Noise is an error and decreases the quality and readability of the peaks. The improvement in SNR is possible because the FTIR spectrometer can easily perform more than one scan per sample: it will simply move its mobile mirror back and forth as many times as the number of scans wanted. The different interferograms are then combined and averaged into one spectrum through a process called signal averaging (SMITH, 1996). Appendix 2 contains a table showing the expected noise reduction of multiple scans compared to one.

4.2.2 Calibration

A crucial element is the calibration process which is instrumental in the derivation and validation of the prediction equations needed in spectrometry. This is done by linking spectral information to the corresponding reference values (GENGLER *et al.*, 2016). The statistical tools used are developed and applied in a specific scientific field called chemometrics. This chemical discipline can be defined as combining mathematics and statistics in order to design or select optimal experimental procedures and computational methods to extract a maximum of relevant chemical information from chemical data (i.e. spectra) (CRA-W, 2017). There are many aspects involved in the calibration process, including the choice of the best reference samples. In the following chapter we will focus on two other issues of large importance, the pre-treatment of the spectra and the multivariate methods used.

4.2.2.1. Pre-treatments

By definition, combining highly correlated spectral data is a difficult endeavour. It was shown that pre-treatments of MIR data can improve the linear relationship between the spectra and the reference values. This makes them an important step to obtain robust prediction models and they are therefore very commonly used (RINNAN *et al.*, 2009). The aim of the two most common types of pre-treatments is to either correct scattering, or improve the resolution of the spectra. Some examples of scatter corrections are multiplicative scatter correction, standard normal variate, and normalization. The most commonly used method of derivation is the one developed by Savitsky-Golay. As impact of pretreatment is not predictable, often calibration models are built with or without (Soyeurt et al. 2011).

4.2.2.2. Multivariate calibration methods

Multivariate calibration methods used in chemometrics can be classified into different types of methods, as seen in Figure 8. 'Multivariate analysis' means that several measurements are used, e.g. several spectra, as opposed 'Univariate analysis' that only uses one measurement, e.g. a particular peak at a time.

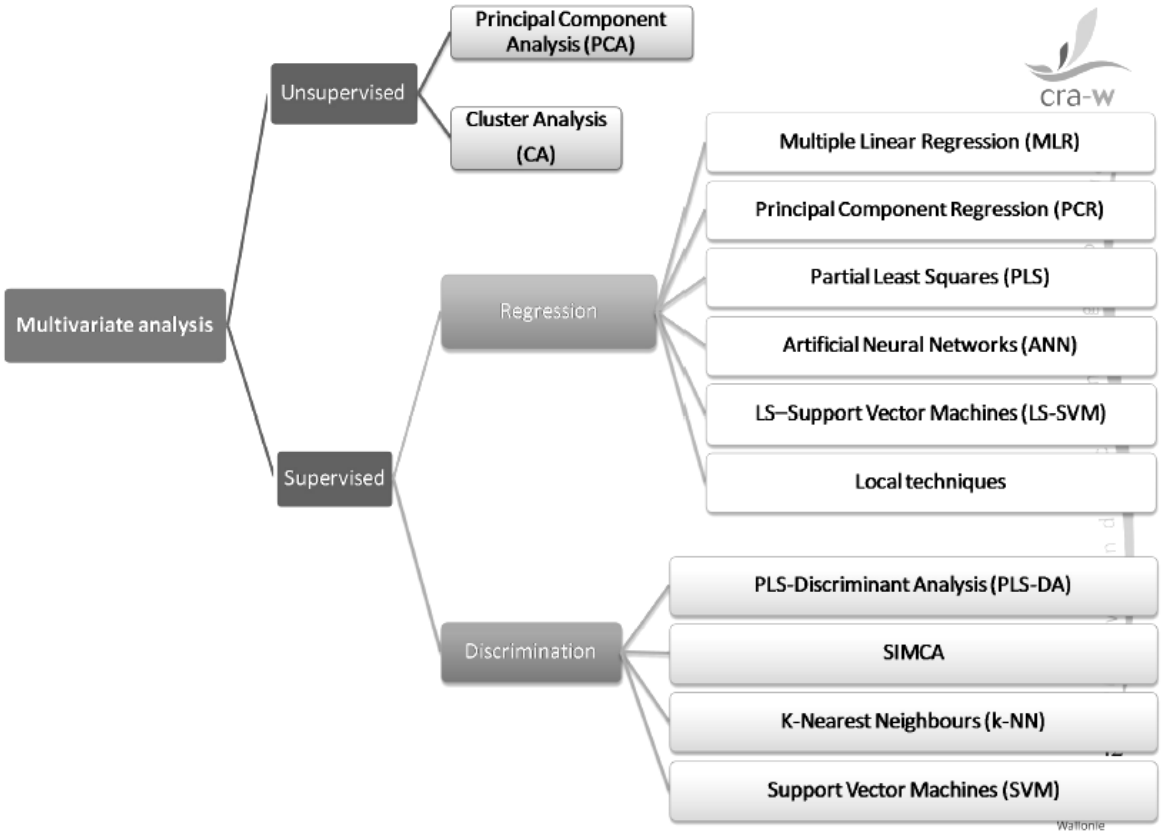

Figure 8: Classification of multivariate methods used in chemometrics (CRA-W 2017).

'Unsupervised' methods only have access to the spectral information, they try to extract patterns without knowing the target, in contrary to 'supervised' methods who rely on spectral information as well as an *a priori* knowledge of the data. Examples of an *a priori*

knowledge are the previous classification into categories or the access to reference values associated with the spectra. In general, this implies the knowledge of the calibration target.

Those falling under the term 'supervised' are model construction methods. They relate reference values or classes to the different absorbance or transmittance values associated to each wavenumber of spectra using a linear equation (CRA-W, 2017; GENGLER *et al.*, 2016). This linear equation is then used in a regression or classification context.

Regression analysis is used when the variable we want to predict is quantitative, e.g. BHB. After calibration, the developed linear equation can then be used for new samples to determine the concentration in which this molecule is present. As explained previously, this is an extension of the Beer-Lambert law. Figure 8 gives the most usual regression methods used, Partial Least Squares (PLS) regression being often favoured. It is close principal component analysis for each group of variables X and Y (if there are more than one variable to predict jointly, in our situation it is mostly only one variable called hereafter Y) but trying to find the multidimensional directions in the X space that explain Y. This implies the projection of X (and Y) to latent structures. This strategy combines the X variable (for MIR in milk up to 1060 data points) into a number of latent variables. Latent variables do not have a tangible counterpart, but they enable us to only have about 10 variables in our model, as opposed to 1060, leading to targeted reduction of variables while explaining most of the variability of Y encountered in the reference sample population. The residual mean square error of the cross-validation (RMSECV) is a good indicator to choose the best number of latent variables. It is best to choose the model with the smallest possible RMSECV).

Not all phenotypes are qualitative. Good examples are classifications into groups (e.g. 'lame or not' or a lameness score of 1 to 5). In this case a classification analysis is required allowing discriminations (CRA-W, 2017). One of these methods is called PLS-discriminant analysis or PLS-DA. It is based on the same principles as PLS, but instead of a regression of a continuous variable, the result is a classification into different classes.

4.3. Milk MIR analysis of milk

The use of MIR spectrometry has become common in dairy cattle, both for milk component testing of individual cows and for milk payment of whole herds, and this for the main components, fat, protein but very often also urea and lactose. Most laboratories no longer use monochromator based spectrometers but FTIR based that measure the spectrum at a great number of different wavenumbers (e.g. 899 for Bentley, 935 for Delta, and 1,060 for Foss instruments (SMITH, 1996; GRELET *et al.*, 2012)). This allows the generation of the whole spectra for further research and development.

4.3.1. Methodology

Figure 9 shows the main steps of the MIR analysis of milk. Milk samples for MIR analysis can be taken during milk routine recording. The sample is either from the morning milk, evening milk or a mixture of both. To be complete, it has to be stressed that milk samples may come from other sources. For example, in order to allow component payments, tank milk samples are also taken on farm by the dairy plants and analyzed using the same

procedure. But we will focus on performance recording samples. These samples are then analyzed by an FTIR spectrometer (FTIR) to produce corresponding MIR spectra.

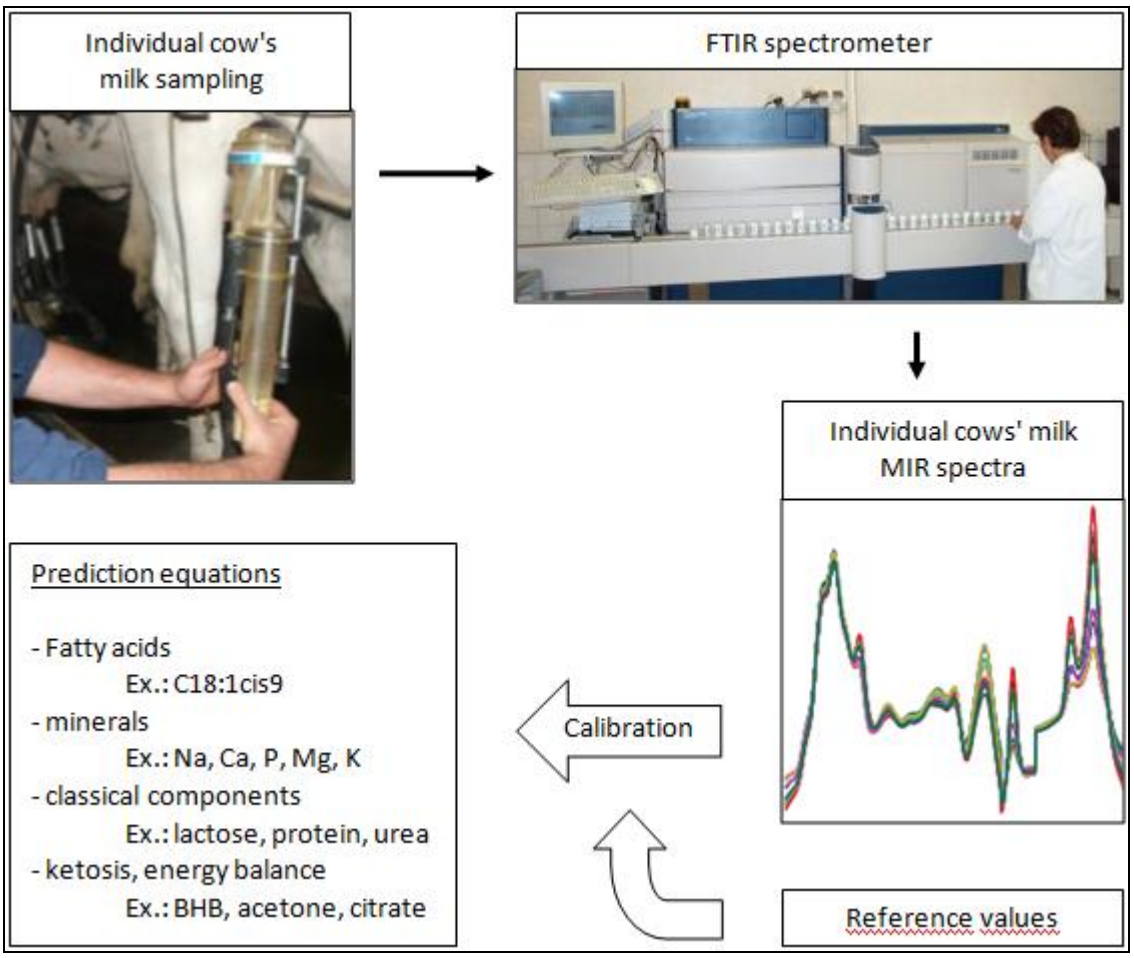

Figure 9: MIR spectroscopy (GRELET et al., 2012)

Until recently MIR spectra were not stored or extracted, only predicted values as fat, protein but very often also urea and lactose were provided to the customers. Recently, new prediction equations were developed.

Additionally reference values for the trait of interest are obtained from the same samples or individuals through reference analysis. Examples of reference analysis are: the Kjeldahl method applied on the milk samples in the case of organic nitrogen, or the locomotion scoring of the cows in the case of lameness. The next step is called calibration and combines the MIR spectra and the reference values using chemometrics to extract the relevant information. This produces prediction equations that can then be used for management or breeding purposes.

4.3.2. MIR spectrum of milk

Figure 10 shows an example of a MIR spectrum of milk.

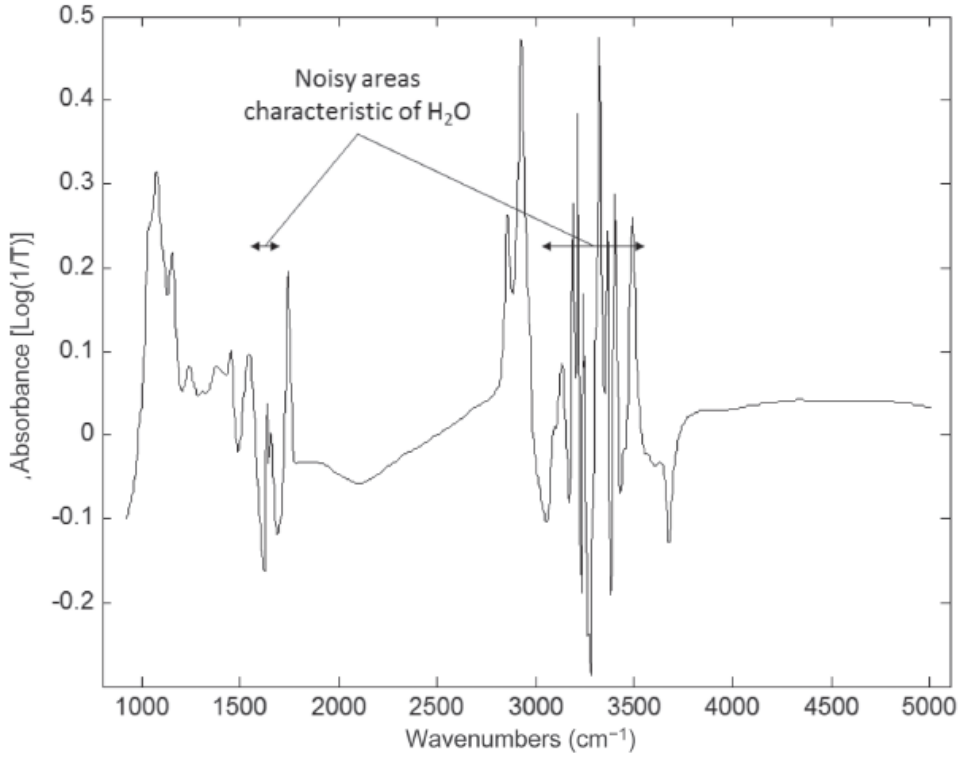

Figure 10: Raw milk spectrum. T=transmittance (GRELET et al. 2015).

The absorption values give information about the composition of the milk as each combination of atoms and chemical bonds absorbs light at a precise wavenumber. Yet, not all those 1060 data points are used in the making of the prediction model. Indeed, some parts of the spectra are 'noisy' because of strong water absorption. Therefore specific parts, containing the most information whilst also reducing the noise to a minimum are selected. These spectral areas, in the case of milk, are: 968.1 to 1577.5 cm^{-1}, 1731.8 to 1762.6 cm^{-1}, 1781.9 to 1808.9 cm^{-1} and 2831.0 to 2966.0 cm^{-1} (GRELET, 2016).

The height or area of a peak is also important because it is proportional to concentration. Spectroscopy is therefore not only a qualitative method that can be used to indicate the presence of a substance, but also as the quantity in which it is present. This link between absorption and concentration is called Beer–Lambert law. In chapter 4.4. we will show how through a generalization of Beer–Lambert law that quantification can be done by combining linearly results at different wavenumbers throughout the considered spectral ranges.

5. Conclusion

A lot of studies have worked on lameness and the detection of it. There is a great interest in the development of fast and accurate ways of lameness detection, especially for early lameness problems, as early treatment reduces the economic impact of the condition and the long-term effects on cow health. There is a link between lameness causes and certain blood and milk components of the cows. This directly links MIR spectra to the condition of being lame. And yet, until now, no studies were undertaken trying to use MIR spectral data

directly or even indirectly, through MIR based predictions of milk components of interest, to assess risks for claw health.

According to ZOTTL *et al.* (2017), a monthly lameness assessment is very promising as a management tool. One locomotion score per month seems enough for early detection, even before drastic changes in milk composition or BCS occur. However, a monthly milk prediction with milk recording samples that are sampled anyway for routine milk control, could be a much more practical way of implementing routine locomotion scoring of dairy cows, provided that the changes in milk, even if not dramatic, are strong enough for early detection. As opposed to having trained staff, from the milk recording association for example, travelling once per month to each farm for locomotion scoring. DE MARCHI *et al.*(2014) mentions the growing interest in methods that routinely and accurately measure and predict animal characteristics such as phenotypes. In this context, MIR technology is very interesting as it is fast, does not require any supplementary workload for the farmer, and is already routinely applied by milk control agencies for evaluations of fat and milk content among other things.

Ultimately, if such an equation can be established, it could be useful both as a routine management tool for farmers, and as a potential help in the selection of cows with a better natural predisposition against the development of lameness. The objective of this work is to assess the feasibility of such a detection or prediction equation of the lameness status of a cow, through the use of MIR calibration.

CHAPTER III: MATERIALS AND METHODS

1. Introduction

This work focuses on the use of MIR for the phenotypic detection of lameness. The aim is to produce a prediction equation that can indicate when a cow is lame, i.e. detection, or if a cow shows signs of future lameness problems, i.e. prediction. If such a prediction equation can be found, it would be an interesting tool in the management of dairy cow health. This work is divided into two main studies. The first study uses classic MIR calibration, while the second is centred around oriented MIR calibration. The following chapter will explain how the data used in this work was sampled, selected, divided into subsets and pre-treated for use in these calibrations.

2. Data sampling and selection

The data for this research was provided by RINDERZUCHT AUSTRIA (2017), from their "Efficient Cow" project. This project was launched in 2012, and collected data of 167 Austrian farms about 5500 dairy cows of the breeds Simmental, Brown Swiss and Holstein. The "Efficient Cow" project was started because of a recent demand by more than 75% of Austrian dairy farmers for further developments in the areas of metabolism, claw health and feed efficiency (RINDERZUCHT AUSTRIA 2017b). In order to do this, many phenotypes were recorded on farms, and through milk recordings (ZOTTL *et al.*, 2017). Among the data collected, there was information on animal health, such as locomotion scores and claw care. Locomotion was assessed by trained staff from the milk recording organization using Visual Locomotion Scoring by SPRECHER *et al.* (1997), described in the literature review and which gives a cow a score from 1 to 5. For the present analysis, an additional classification variable was created out of these scores, classifying animals into non lame and lame, like Table 3 shows. Two was picked as a threshold, as this slight change in gait, described in the 5-point locomotion scoring system, could also be attributed to mild stiffness, for example when standing up after lying down for a certain amount of time. Because all locomotion scoring staff was trained in using the same technique, it can be assumed that the results from different staff members were equivalent. Diseases or problems of the foot or hoof were recorded by hoof trimmers (ZOTTL *et al.*, 2017).

Table 3: Threshold classification of the cows into lame or none lame, based on their locomotion score. Locomotion score, and description by SPRECHER *et al.* (1997).

Locomotion score	Description	Lameness classification
1	Normal	No
2	Mildly lame	No
3	Moderately lame	Yes
4	Lame	Yes
5	Severely lame	Yes

The milk samples used for MIR analysis were gathered during routine milk recording and were analyzed by FTIR-MIR spectrometry using 1 FOSS FT+ and 2 FOSS FT6000 instruments. The MIR instruments were standardized using the EMR/CRA-W standardization process (GRELET *et al.*, 2015). This system of standardization matches spectra from different slave-instruments to one master-instrument spectra. This enables the combination of spectra coming from different instruments into one common database which can then be used, as if they were from the same instrument. The MIR data were recorded between July 2014, start of the standardization of the spectrometers in Austria, and December 2014 on 3895 cows from 122 farms, giving a total of 9324 records. Cows were recorded for spectral data an average 2.4 times, and the obtained standardized spectra were put in absorbance.

Table 4: Data collected by the Efficient Cow Project and provided by RINDERZUCHT AUSTRIA.

Data set	Variables
Phenotype data	Farm, animal, breed, locomotion score, date of locomotion scoring, culling reason, parity, pregnancy
Spectral data	Standardized absorbance spectra, date of milk recording
Predicted values	Milk ketosis molecules, major components, fatty acids, minerals
Housing data	Tie or free stall housing, barn floor surface, bedding type
Pasture data	Access to pasture
Hoof trimmer data	Affected claw, type of lesion, severity, date of hoof trimming
Veterinarian data	Type of lesion, date of diagnosis

Table 4 shows the data shared by RINDERZUCHT AUSTRIA through different data sets. Additional MIR predicted traits were provided by Dr. Clément GRELET using prediction equations established by the CRA-W, GxABT and collaborators on the spectral data from Austria. These traits were the concentrations in fatty acids, in components linked to ketosis and energy balance, in minerals and in some major milk components. In addition to predicted traits, the standardized Mahalanobis distance (GH) was computed from each spectrum and is the distance to the central point of the Fatty acids database. It enabled the deletion of outlying records which are not covered by the variability of the model. The diagram in Figure 11 gives a visual representation of the steps taken during the combining of data sets, and the selection of relevant data. Data preparation was done using procedures included in the statistical suite SAS (SAS INSTITUTE INC., 2017), and the presentations of results in tables and graphs were done in Microsoft Excel (2007).

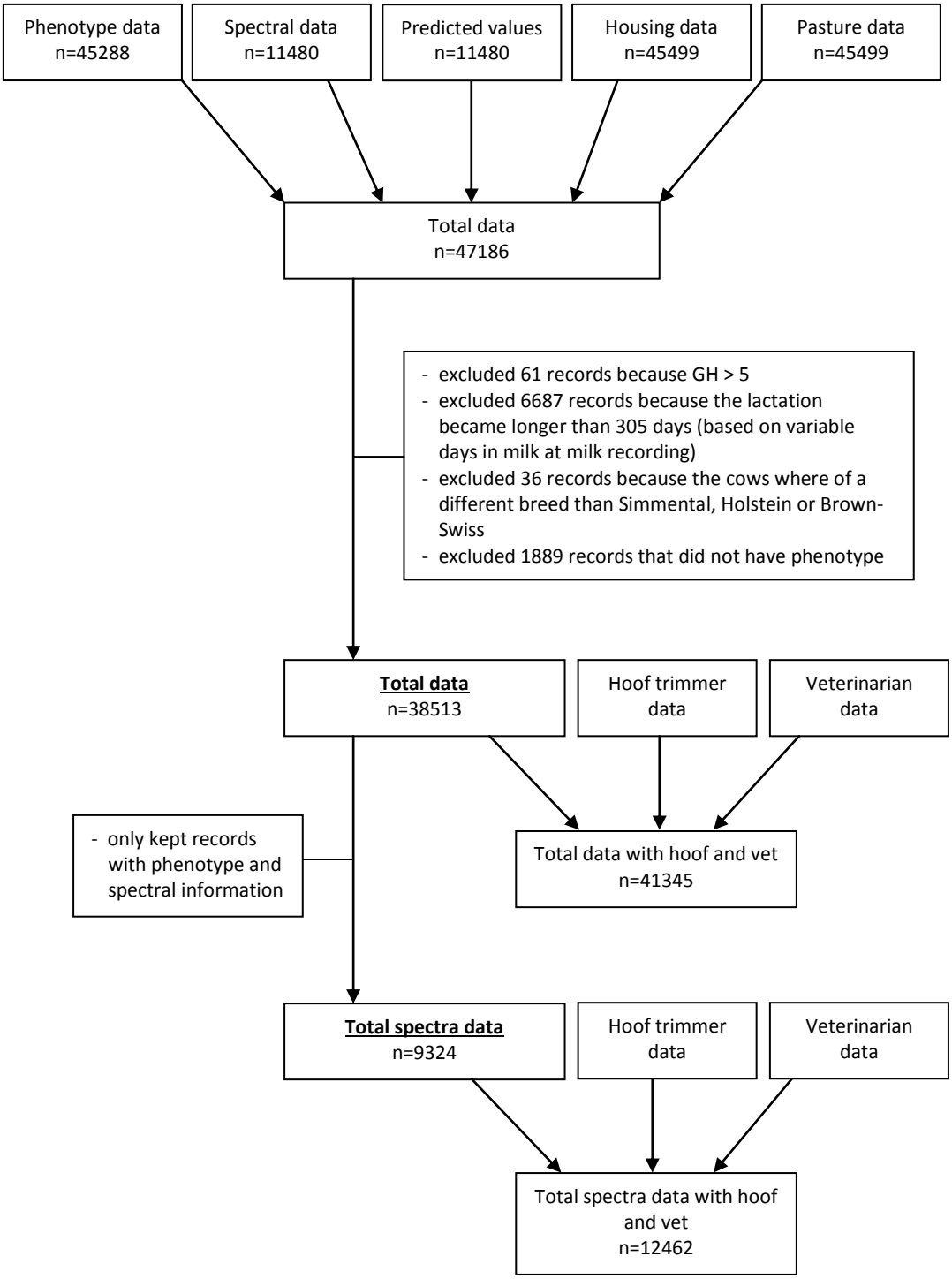

Figure 11: Combination of data sets and elimination of outliers.

Red Holsteins and Black Holsteins, as being genetically very close populations, were combined into one group called Holstein. In this study, international names for breeds were preferred. The Brown Swiss of this study, even if they are called Brown Swiss in German, are indeed Brown Swiss cows, and not Original Brown Swiss animals without North-American ancestors. The Simmentals used in the study were not Swiss Simmental, but Austrian Simmental, which originated from cross-breeding of local stock with Swiss Simmental in

Bavaria and Austria in the 19th century. Other minor breeds present in the original data sets, were deleted.

The data set used for most of the calibration work is the one called 'Total spectra data' in the above diagram. It contains records of a lactation duration limited to 305 days with locomotion scoring, and spectral information for all records. The data set 'Total data' contained phenotypic information, and therefore locomotion scores for all animals, but no spectral information for all animals. This data set was used for descriptive statistics purposes.

The following tables give a rough overview of the two most important data sets. The first gives a good idea of the general structure of the farms. We can see that most of the animals were of the Simmental breed, with the number of Brown Swiss and Holstein cows being lower at about 1500 and 1200 respectively. The mean number of animals per breed does not give a lot of information, because 65 of the 161 farms actually have mixed herds, with some of their animals being of a different breed than the majority of the herd. This minority group can go from 1 cow to almost half the herd.

Total data		Records	Animals	Farms	Cows per farm			Records per animal		
					Mean (st d)	Min	Max	Mean (st d)	Min	Max
	Simmental	22 729	3756	113	33.2 (19.5)	1	93	6 (2.5)	1	12
	Brown Swiss	8 896	1475	56	26.3 (22.5)	1	111	6 (2.6)	1	12
	Holstein	6 888	1203	65	18.5 (25.5)	1	115	5.7 (2.6)	1	12
	Total	38 513	6434	161	39.9 (21.1)	3	116	6 (2.5)	1	12

Total spectra data		Records	Animals	Farms	Cows per farm			Records per animal		
					Mean (st d)	Min	Max	Mean (st d)	Min	Max
	Simmental	6428	2597	96	27.1 (14.1)	1	74	2.5 (1.1)	1	5
	Brown Swiss	1393	646	25	25.8 (19.6)	1	88	2.2 (1.3)	1	6
	Holstein	1503	652	43	15.2 (22.4)	1	89	2.3 (1.1)	1	6
	Total	9324	3895	122	31.9 (16.5)	7	92	2.4 (1.1)	1	6

A 2012 report by the European Commission (EUROPEAN COMMISSION, 2012) gives an average herd size of 20 dairy cows for Austrian dairy farms in 2009. The average farm size of the present data is 40 animals, but with great variability, like shown in the above table, with herd sizes going from 3 to 116 animals. Because the farmers took part on a voluntary basis, there was a small, automatic selection of the 'bigger' and 'best' farms in Austria, which explains the larger herd sizes. However, we are trying to construct an 'individual cow' model and in that case, the data is still representative enough of the Austrian dairy sector.

Table 5: Size of the farms depending on the number of dairy cows per herd.

Herdsize (number of dairy cows)	Number of farms
3 to 20	20
21 to 40	70
41 to 60	45
61 to 80	15
81 to 100	4
101 to 116	4

When hoof trimmer information was needed, two other data sets were used. These were 'Total hoofvet data' for descriptive statistics, and 'Total spectra hoofvet' data for MIR calibration. They were only used when really needed, as adding the hoof trimmer information sometimes created duplicates of cows for the same record, because more than one claw of the same cow had a lesion. This way, having duplicates of spectra for calibration was reduced to a minimum.

3. Interactions of lameness with environmental and individual factors

3.1. Factors that influence the prevalence of lameness

For this part, different aspects of the records, such as breed, parity and access to pasture, were looked at and brought into relationship with the locomotion scores and registered claw lesions given for those records. SAS software (SAS INSTITUTE INC., 2017) was used to select the data, perform Pearson chi-square tests and find Pearson correlation coefficients for ordinal variables like parity.

3.2. Milk components potentially linked to lameness

3.2.1 Looking at specific animals

The values predicted by Dr. Clément GRELET, from the spectra provided by RINDERZUCHT AUSTRIA, make it possible to look at the variations in milk composition for sound and lame animals. The values that were predicted, and that could be used as potential biomarkers for lameness, are the following:

- Major components: protein content
- Fat: fat content, fatty acids like C18:1cis9, short chain, mid chain and long chain fatty acids
- Milk ketosis molecules: BHB, acetone, citrate
- Minerals: Ca, Na, P and K

These traits will from hereon be referenced to as lameness biomarkers, or in short biomarkers in the rest of this work.

To get an idea of the evolution of the locomotion score and the biomarkers over the duration of a lactation, certain cows were singled out and studied in more detail. Cows chosen for this, had at least 5 locomotion scorings over the course of their lactation, and spectral data for the corresponding months. They also needed to have been given a locomotion score of 5, at least once, to make sure that they would show a wide variety of locomotion scores over the duration of their lactation. The results for two of these cows are shown in results and discussion in the form of graphs, and tables with the values for each biomarker can be found in Appendix 1.

Table 6: Descripion of the observed cows.

	Breed	Parity	Number of records	Lameness scores
Cow 1	Simmental (FL)	5	5	3, 3, 2, 5, 4
Cow 2	Brown Swiss (BV)	5	5	1, 2, 4, 3, 5

To be able to compare the increase or decrease of the different biomarker in the same graphs more easily, the values of certain traits had to be divided or multiplied by multiples of 10, so that all values are between 0 and 2. BHB was divided by 100, citrate by 10, Ca, Na, P and K by 1000 and acetone was multiplied by 10. This improves the quality of the graphs and shows variations better.

3.2.1 Generalization

The aim of this part was to see if the findings obtained by looking at specific cows, could be generalized to the whole sample. To this end, the values of biomarkers between cows of locomotion score 1 and locomotion score 5 were compared. Additionally, the Pearson correlation coefficients were computed between locomotion score, and the different biomarkers.

4. First study: Classic MIR calibration

In classic MIR calibration, latent variables are extracted from the X data, i.e. spectral milk information, in such a way that they explain as much as possible the overall variation encountered in this X data. In this study, we chose to reduce the variation due to known factors, by making data sets smaller, and by reducing the variation in the spectra. This allowed obtaining better results for the classification of records into lame or non lame.

This study is divided into two parts. The first part and the corresponding results are part of an article submitted for review to the "Agriculturae Conspectus Scientificus" about classical MIR calibration for lameness, in the framework of the Animal Science Days (ASD), an international symposium taking place in September 2017. After submission of this article, several alternative and improved strategies leading to novel results, were developed and used. They are the object of the second part of this study.

4.1. Animal Science Days article (adapted from Mineur *et al.*, 2017)

4.1.1. Data selection

The data used in the calibration models presented in the article submitted for review, differs slightly from the general data described in the descriptive statistics, and used in the other calibration models as several edits were not done yet, when this study was submitted. Here follows a list of differences between the 'Full ASD data' set, 9811 records, and the 'Total spectra data' data set, presented in the first part of this chapter, and containing 9324 records:

- The maximum duration of lactation was 365 days. This was changed to 305 days in the additional computations (4.2.), because 305 days is used as default lactation length for cows. Using this limit, is more logical than using one year.

- The maximum number of days between the locomotion scoring and the milk recording, was 7 days. This was changed as poor metabolic status can have long-lasting effects on lameness, and hoof lesions can take time to heal and have an influence on the cow's condition for more than 7 days. It therefore seemed logical to abandon this restriction. With no limitation set, the longest separation was 19 days.

- At the time, when the 'ASD' calibrations were realised, the predicted values had not been predicted yet, including the Mahalonobis distance. Therefore no GH based quality checks could be done on the spectra.

These are the reasons for differences encountered between both complete data sets, i.e. 'Full ASD data' and 'Total spectra data'.

4.1.1.1. Subsetting of the full asd data set

Different data sets were created. The full ASD set (9811 records) was used as a reference. Multiple sub sets, either linked to a specific period in the lactation, to specific diseases, to a breed or to parity, were created to obtain more homogeneous data sets. This was done to see if homogeneity of the dependent variable definition, and the associated spectra, had an influence on the precision of the model. Subsets with different groupings of animals were created with only heifers (parity = 1), young (parity = 1 or 2) and old (parity > 2) cows. This was done as there is not a consensus on the influence of parity on milk composition in the literature. Some studies report this influence (e.g. YADAV *et al.*, 2013; YANG *et al.*, 2013) others do not (e.g. GURMESSA *et al.*, 2012). Subsets were also created for lactation stages; a factor that is known to influence milk composition (e.g. BASTIN 2011). In order to allow differences, due to health and lameness, having a better chance of standing out, we also took breed into account to smooth out breed related differences. Indeed, many studies (e.g. HEINRICHS *et al.*, 1997) reported the influence of breed on milk composition. Heel horn erosion (HHE) is the dissolution and decay of the horn on the bulbs of the heel. White line

disease (WL) refers to a gap between the sole, and the wall often filled with faeces or decayed horn masses, which can lead to an abscess if the leather skin is affected (EGGER-DANNER *et al.*, 2015). For both diseases, specific sets were created by selecting all records with the disease, adding all the records of healthy animals coming from the same farms, and by only keeping records were the hoof trimmer data had been collected within 3 weeks of the MIR analysis. However, not all animals affected by HHE or WL, were also lame so 2 extra files were created, where only diseased lame and healthy non lame animals were kept.

4.1.1.2. Calibration and validation set selection

A third of the records of every data set, were randomly selected for validation. The other two thirds of each subset, were selected and used for calibration. For this separation of calibration and validation, the data sets were separated by record, not by animal, even though some animals have multiple records, because the lameness status of an animal can vary over time and the animal may find itself in different living conditions, e.g. pasture in summer and stall in winter, which makes every record unique. In the case of heel horn erosion (HHE) and wall defect (WL), some MIR spectra found themselves in both data sets, as the cow had one of her hoofs affected by heel horn erosion, while another had a wall defect problem. In total, from the 9811 records, 1843 had HHE and 1068 suffered from WL disease.

Table 7: Number of records (N), of lame records (L) and number of farms (F) for each calibration, and validation subset of the main dataset 'Full ASD data'.

	Total		Calibration		Validation	
Subset	N	L	N	L	N	L
All	9811	795				
First half of lactation	5509	490	3673	312	1836	178
Last half of lactation	4302	305	2868	203	1434	102
First third of lactation	3806	348	2537	239	1269	109
Last third of lactation	2479	176	1653	111	826	65
Simmental	6828	578	4552	388	2276	190
Holstein	1560	121	1040	77	520	44
Brown Swiss	1423	96	949	69	474	27
Heifer (parity = 1)	2792	96	1861	64	931	32
Young (parity = 1 or 2)	4855	195	3237	133	1618	62
Old (parity > 2)	4956	600	1652	395	1652	205
HHE*, 3 weeks	596	52	397	32	199	20
HHE* & lame, 3 weeks	273	52	182	39	91	13
WL**, 3 weeks	678	41	452	29	226	12
WL** & lame, 3 weeks	465	41	310	26	155	12

4.1.2. Spectral data pre-treatment

Pre-treatment of the spectra, consisted of a first or second derivative with widths of 5, using the Savitzky-Golay method, to enhance resolution, and eliminate additive baseline drift between samples. This was followed by a transformation to Standard Normal Variates (SNV), to standardize each spectrum into having a mean of 0 and standard deviation of 1, to correct scattering (FEARN, 2017; HUANG *et al.*, 2010). Venetian blinds were chosen as cross-validation, which means 10% of the calibration set was randomly selected 10 times.

Yet, not all those 1060 data points are used in the making of the prediction model. Indeed, some parts of the spectra are 'noisy' because of strong water absorption. Therefore specific parts, containing the most information whilst also reducing the noise to a minimum, were selected. These spectral areas are: 968.1 to 1577.5 cm-1, 1731.8 to 1762.6 cm-1, 1781.9 to 1808.9 cm-1 and 2831.0 to 2966.0 cm-1 (GRELET, 2016). It is also important to note that lameness prediction is indirect, as it is not directly measured in milk, but based on modifications in milk composition; therefore the selected parts of the spectrum are used as a whole to predict lameness directly. Illustrations of the effect of data pre-treatment on MIR spectra can be found in appendix 1.

4.1.3. Calibration

Prediction models were done with Partial Least Squares Discriminant Analysis (PLS-DA), using the software PLS-Toolbox, by Eigenvector Research Inc., implemented on the Matlab software (THE MATHWORKS INC., 2000). The PLS-DA is a variant of PLS regression, which is used when the dependent variable is categorical, in this case lame vs. non lame (FERNÁNDEZ PIERNA, 2017). The type of cross-validation used, is called Venetian blinds, and it selects ten times 10% of the calibration data for cross-validation.

The number of latent variables used in the prediction models, was chosen based on the break of slope of the Root-Mean-Square Error of Cross-Validation (RMSEcv) plot of the data set ‘all’. RMSEcv is a measure of fit, and the smaller this value, the better the prediction model. The break of slope is the point where adding another latent variable, does not significantly reduce the RMSEcv anymore.

4.2. Additional computations

4.2.1 Data selection

4.2.1.1 Subsetting of the total data set

The full data set, 'Total spectra data', was first split into two main detection and prediction sets, and only then split further into subsets. The main dataset 'Detection' comprises records for which locomotion scoring happened before or on the same day as MIR analysis, and the main 'Prediction' dataset is made of the records for which MIR analysis happened before locomotion scoring. This was done, because different interactions between biomarkers and lameness are considered in the two directions. Yet, a risk of separating detection and prediction, is that a great majority of records were part of prediction, which considerably reduces the detection data sets.

The subsets that concerned lactation duration, were split differently from 4.1.1.1. because milk yield over lactation is not linear, and the period that is most taxing for the animal is the first two months, when the cow has to recuperate from the calving, start her lactation, and increase milk yield until she reaches a peak after about two months. Therefore, the decision was made to split the subsets into 2 and 8 months, instead of first and last half, or thirds of lactation, which are purely temporal measures that have nothing to do with the production of the cow.

Three new subsets were also tried for the 'Prediction' dataset. The records of the first two subsets, 'Locomotion score 1, 3, 4 and 5' and 'Locomotion score 1, 4 and 5', were selected in such a way that locomotion score 1 and locomotion scores 1 and 2 were missing, respectively. This was done to see if asking the model, to predict more 'extreme' values, would help with the predictions. For the third subset, the number of records with locomotion score 1 was randomly selected to only include 1000 records with locomotion score 1, instead of 4173. This makes it more balanced in comparison to the other locomotion scores: 836 records with score 2, 278 with score 3, 118 with score 4, and 26 with score 5. The aim was to see whether a more 'balanced' set, would help with the prediction.

4.2.1.2 Calibration and validation set selection

The validation and calibration data sets were still split randomly into two thirds and one third, but according to farms instead of records. This was done to reduce overfitting because it prevents records of the same cow to end up in both calibration and validation sets at the same time. Like was explained earlier, every record can be considered unique, because conditions of the cow change over time. However, the farm in which a cow lives, forms an important part of the environment and conditions of a record. So, it was chosen to separate by farm. The number of records, lame records and farms for each dataset is summarized in Tables 8 and 9.

Table 8: Number of records (N), of lame records (L) and number of farms (F) for each calibration, and validation data set of the main dataset 'Detection'.

DETECTION	Total			Calibration			Validation		
Subset	N	L	F	N	L	F	N	L	F
Detection	1993	160	57	1314	98	19	679	62	38
First 2 months of lactation	438	46	57	287	29	19	151	17	38
Last 8 months of lactation	1555	114	57	1127	79	19	428	35	38
Simmental	1604	123	50	1140	78	17	464	45	33
Brown Swiss	184	7	7	154	7	2	30	0	5
Holstein	205	30	14	165	28	5	40	2	9
Heifer (parity = 1)	533	15	57	353	8	38	180	7	19
Young (parity = 1 or 2)	943	35	57	577	19	38	366	16	19
Old (parity > 2)	1050	125	56	599	48	37	451	77	19

Table 9: Number of records (N), of lame records (L) and number of farms (F) for each calibration, and validation data set of the main dataset 'Prediction'.

PREDICTION	Total			Calibration			Validation		
Subset	N	L	F	N	L	F	N	L	F
Prediction	7331	592	115	4736	400	77	2595	192	38
First 2 months of lactation	1526	138	113	1006	83	75	520	55	38
Last 8 months of lactation	5805	454	115	3677	278	77	2128	176	38
Simmental	4824	414	91	3463	294	61	1361	120	30
Brown Swiss	1209	90	23	865	80	15	344	10	8
Holstein	1298	88	41	859	45	27	439	43	14
Heifer (parity = 1)	2117	75	113	1437	54	75	680	21	38
Young (parity = 1 or 2)	3677	145	115	2321	98	77	1356	47	38
Old (parity > 2)	3654	447	114	2375	308	76	1279	139	38
Locomotion scores 1, 3, 4 and 5	6209	592	115	4354	422	77	1855	170	38
Locomotion scores 1, 4 and 5	5821	204	114	3907	133	76	1914	71	38
Balanced locomotion score 1	4158	542	115	2258	422	77	1900	170	38

4.2.2 Spectral data pre-treatment

Results obtained using data and methods from 4.1 (see 3.1.1.) were incorporated into improved strategies in these analyses. Only the first derivative was used, as the second derivative had a tendency to lower the sensitivity in favour of the specificity. However, finding the true positives, i.e. the truly lame animals, is more important as it would point out cows that need treatment. Results (Appendix 3) also confirmed that SNV was not necessary.

SNV is used mostly to reduce errors because of differences in granulometry, i.e. size of the analyzed particles, which is usually not a problem when working with milk. (CRA-W, 2017).

4.2.3 Calibration

Calibration was again done by PLS-DA, using the same programs, for classification of records into lame or non lame, based on a threshold of locomotion score 2.

5. Second study: Oriented MIR calibration

Oriented MIR calibration functions in a similar way to classic MIR calibration, in that it also uses milk spectra to predict the status of a cow. The difference resides in how many times the spectra are used. While classic MIR calibration uses the spectra only once, oriented MIR calibration uses it twice or several times; once as usual spectral data, and once through biomarkers, predicted from those spectra.

The effect this has, is that, instead of blindly combining the spectra into explaining the most possible variation displayed by the X, the model combines the spectra in such a way that they try to explain the most possible variation encountered by these predicted biomarkers. So, if the biomarkers have a close link to the Y, this directs the prediction into explaining the Y. So, in a way, the added biomarker or biomarkers force the 'explanation of variability' by the spectra in a particular direction, and the latent variables are selected to explain as much as possible in this direction.

In the case of lameness, the chosen biomarkers are the ones described in 3.2.1. They are BHB, acetone, citrate, F:P ratio, C18:1cis9, LCFA, MCFA, SCFA and Ca. Ca was chosen among the 4 minerals because it is the one which is most intimately linked to hoof horn quality (TOMLINSON *et al.*, 2004), and because of the link with hypocalcaemia. Using these in the calibration, indirectly orients the model towards predicting lameness though the link between these biomarkers, and lameness.

5.1 Data selection

For the oriented MIR calibration, only the main 'Prediction' set from title 4. was used because it has a larger number of records. From this data set, a calibration and a validation set were selected by farm. The size of the obtained sets is summarize in table 10.

Table 10: Number of records (N), of lame records (L) and number of farms (F) for the oriented MIR calibration and validation data sets of the main dataset 'Prediction'.

	N	L	F
Prediction	7331	592	115
Calibration	4736	400	77
Validation	2595	192	38

5.2 Spectral data pre-treatment

As in the other study, first derivative was used to enhance resolution. For this study however, it was then followed by an 'autoscale' pre-treatment. This is because the biomarkers and spectra are sometimes of a very different scale, and autoscale enables to still have them in the same model by centring the spectra around zero, and scaling to unit variance (CRA-W, 2017).

5.3 Calibration

The calibration was once more performed with PLS-DA, with the notable distinction that the PLS-DA extracted latent variables combined in such a way that they are steered towards the biomarkers in the model, and therefore indirectly to towards lameness.

CHAPTER IV: RESULTS AND DISCUSSION

1. Introduction

The first part of the results and discussion is focused on the study of the present data, in particular the factors that influence the prevalence of lameness and the different indicators of lameness problems in the milk composition. Afterwards, this knowledge is used in the elaboration of MIR calibration models for lameness, at first, in classic MIR calibration and then is oriented MIR calibration using MIR based biomarkers.

2. Interactions of lameness with environmental and individual factors

2.1. Factors that influence the prevalence of lameness

For the present data, the general prevalence of lameness was 23.16%, and that of moderate to severe lameness was 8.07%. The prevalence of lameness can be influenced by a host of factors, ranging from cow specific factors like breed or parity, to management factors like housing or access to pasture. Some of these factors are discussed in this chapter.

Table 11: Frequencies in number of records and prevalence in percentages of the different locomotion score.

Lameness category Locomotion score	Sound 1	Mildly lame 2	Moderately lame 3	Lame 4	Severely lame 5	Total
Simmental	5028	863	343	169	25	6428
Brown Swiss	1091	205	67	22	8	1393
Holstein	1046	339	87	27	4	1503
Total	7165	1407	497	218	37	9324
%	76.84	15.09	5.33	2.34	0.4	100

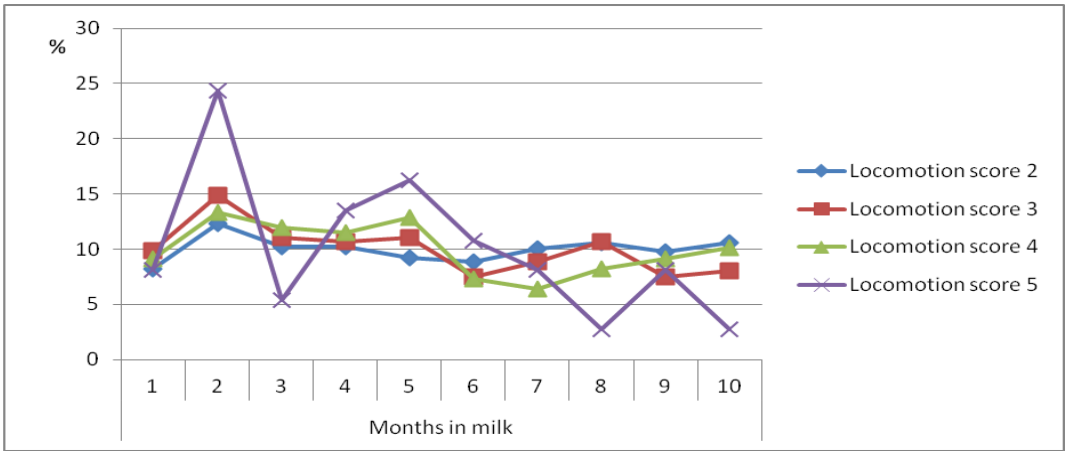

Figure 12: Repartition, in percentages, of the locomotion scores over the duration of the lactation, divided into months in milk.

The repartition of the locomotion scores over the duration of the lactation is represented in Figure 12. It is hard to draw conclusions for locomotion score 5 as only 37 animals were scored that high. The other scores seem pretty level across the lactation with an increase of about 5% during the second month. This peak of lameness during the peak of lactation is linked to the great energy mobilization that takes place at the start of the lactation and has repercussions in the following months (METZNER *et al.*, 1993; COLLARD *et al.*, 2000).

2.1.1. Feet and leg disorders

Not all records contained hoof trimmer or veterinary information, but 26 918 records did. Of those, 49.6% had no claw or leg disorder, and 2.93% had an undefined problem. Table 12 shows, for each type of lesion encountered, the frequency and percentage relative to the total amount of lesions.

Table 12: Frequency and prevalence of different lesions.

Veterinary or hoof trimmer code	Frequency	% of lesions	Description
BF	3365	26.34	Heel horn erosion
WD	2461	19.26	White line defect (fissure or abscess)
DD	1896	14.84	Dermatitis digitalis
SG	1700	13.31	Sole ulcer
DS	1017	7.96	Double sole
SB	700	5.48	Sole haemorrhage
LI	461	3.61	Limax
62	294	2.30	Claw ulcer
SW	195	1.53	Phlegmon
KR	145	1.14	Laminitis
RK	141	1.10	Corkscrew claw
KV	111	0.87	Concave dorsal wall
64	98	0.77	Fractures, dislocations or other leg injury
67	77	0.60	Joint swelling
63	27	0.21	Claw fissure
68	26	0.20	Stuck lying down because of a problem with the locomotion system
65	25	0.20	Muscle or tendon injuries
66	22	0.17	Paralysis
SK	14	0.11	Splitting claws
Total	12 775	100	

The most common problems were heel horn erosion, white line disorders, dermatitis digitalis, sole ulcers and haemorrhages, double soles and limax. Together, they represented about 90% of the lesions found. Next to those, a whole array of less frequent problems was encountered. Some problems did not concern the hoofs, but other parts of the legs like the bones, muscles or tendons of the animal, or were of a different order like paralysis or not being able to stand up again. In the end however, most problems were linked to the feet.

Of the most common problems, the one most likely to provoke strong lameness was limax (LI) with almost 20% of cows being moderately to severely lame as seen in Figure 13. Only heel horn erosion (BF) caused less than 10% of animals to develop moderate to severe lameness. Most disorders caused about 30 to 40% mild to severe lameness.

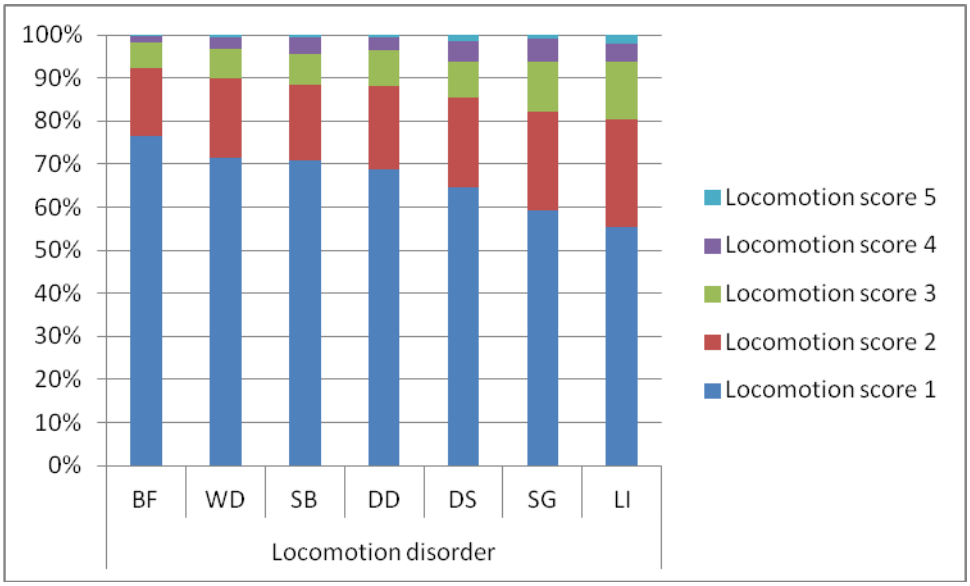

Figure 13: Locomotion scores in function of hoof lesion. Lesion codes from Table 12: BF is heel horn erosion, WD is white line defect, SB issole haemorrhage, DD is dermatitis digitalis, DS is double sole, SG is sole ulcer and LI is limax.

However, this figure also shows us that not all lesions cause lameness. This was already brought forward in the literature review: the conclusion was that the degree of severity of the lesion and the pain linked to it, play an important role on the development of lameness. Only 3741 records came from cows that had been examined by a veterinarian and therefore had been given a severity degree of 1 to 3 to their lesion. Figure 14 illustrates this relationship between lesion severity and lameness: a correlation of 0.16 was found between locomotion scoring and lesion severity. Even then however, we can see that lesions judged to be severe (3) by a veterinarian only caused a little more than 40% of animals to become lame. This again illustrates the complexity of lameness. It may also render classification through milk composition more difficult as animals, which may have milk composition changes due to a lesion, may still walk soundly.

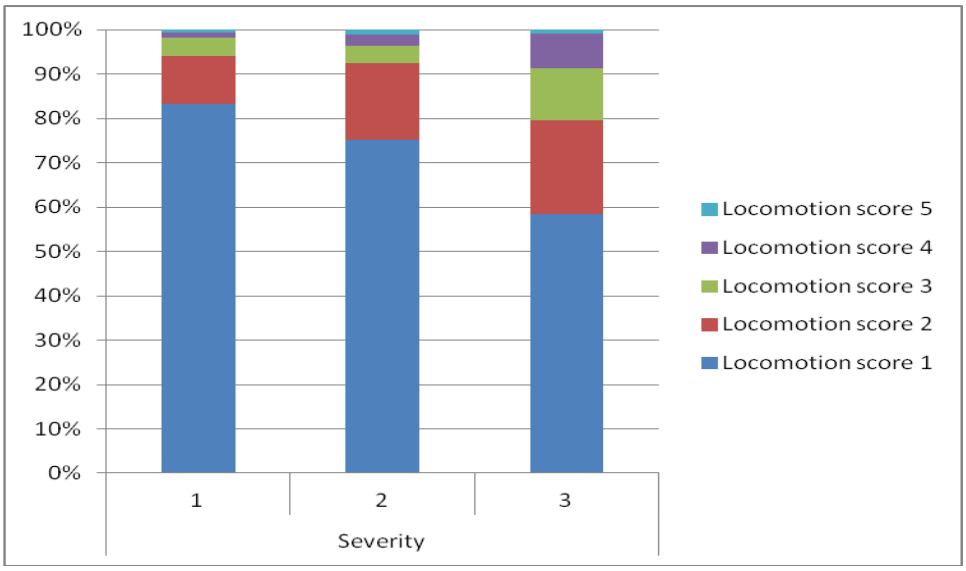

Figure 14: Illustration of the repercussion of the severity of a lesion on locomotion scores.

During the literature review, it was mentioned that most lesions happen to the lateral hind claws. This was also the case for the present data. Figure 15 shows that more than 45% of lesions were found on the left (L) and right (R) outer (AU) claws. The inner (INN) claws seemed equally sensitive to lesions on front and hind feet. However, the space between the claws (SP) again seemed much more prone to lesions on the hind than on the front feet.

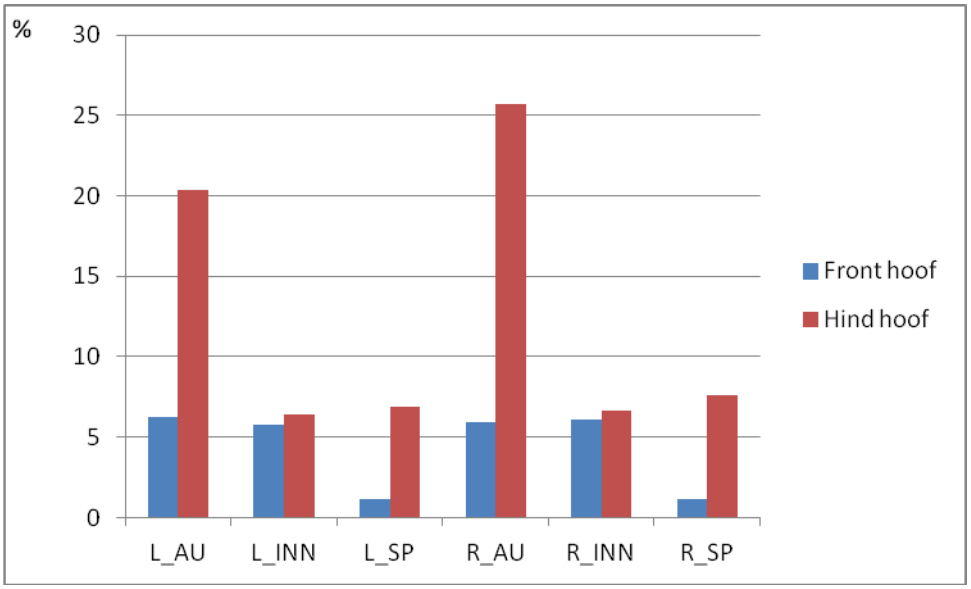

Figure 15: Percentages of lesions found on the different claws and inter-claw spaces of the front and hind feet.

2.1.2. Breed

Figure 16 illustrates how the breed of a cow has an important impact on the development of lameness. This was confirmed by the p-value (<0.0001) of a chi-square test that indicated a very highly significant association between both traits. Holstein (HO) clearly have the most problems with lameness. Simmental (FL), on the other hand, obtain more sound records. On their website (RINDERZUCHT AUSTRIA, 2017D), RINDERZUCHT AUSTRIA gives a description of the most important cow breeds in Austria. Simmental, Brown Swiss (BS) and Holstein are the three most important dairy breeds. Even if lameness prevalence is highly influenced by other factors like age, housing and diet, looking at the characteristics of these breeds can help explain some differences between them.

Simmental is a dual-purpose breed which usually gives the breed a somewhat better BCS than Holstein and puts it less at risk of metabolic disorders, due to loss of condition, that can cause lameness (RINDERZUCHT AUSTRIA, 2017d). DIPPEL *et al.* (2009) also point out that being a dual-purpose breed could reduce the risk of lameness due to over-conditioning at drying off, because the claws are prepared to carry higher loads as well. Lower milk yields might also play a role in the reduction of lameness problems as the Simmental usually produces about 6000 litres of milk in her first lactation and 7000 to 9000 in the following ones. In comparison, the Holstein produces up to 10000 litres per lactation (RINDERZUCHT AUSTRIA, 2017c). High milk yield is a risk for lameness, especially at the start of the lactation (BICALHO *et al.*, 2009).

Even if the original Brown Swiss was a dual-purpose breed, the Brown Swiss (BS), that was selected from it, is a dairy breed with milk yields up to 9000 litres. However, they are characterized by strong hoofs and good ankles, which could explain the smaller number of lesions that occur to their claws, like shown on Figure 17 (RINDERZUCHT AUSTRIA, 2017a).

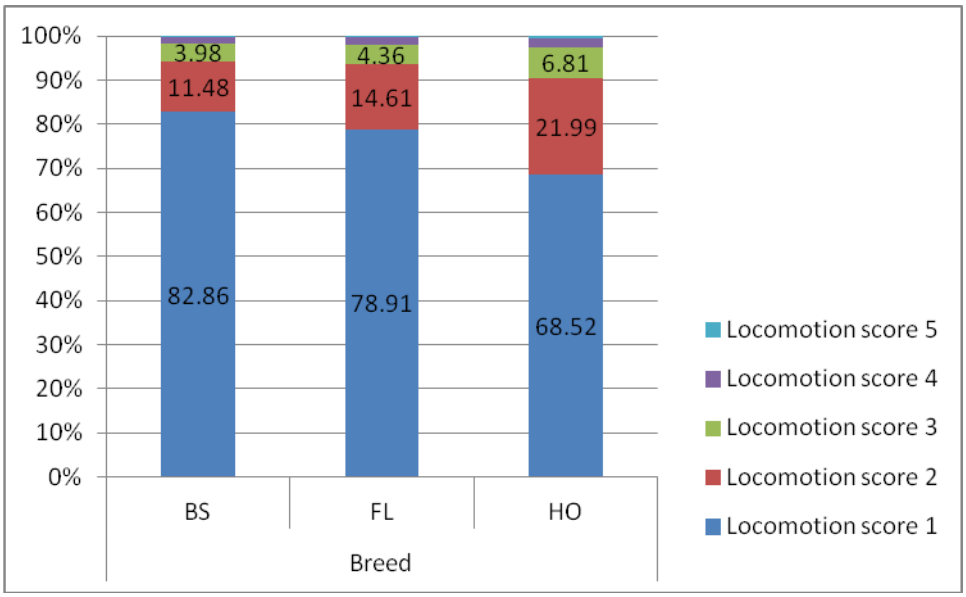

Figure 16: Percentage of records with a certain locomotion score in function of cow breed.

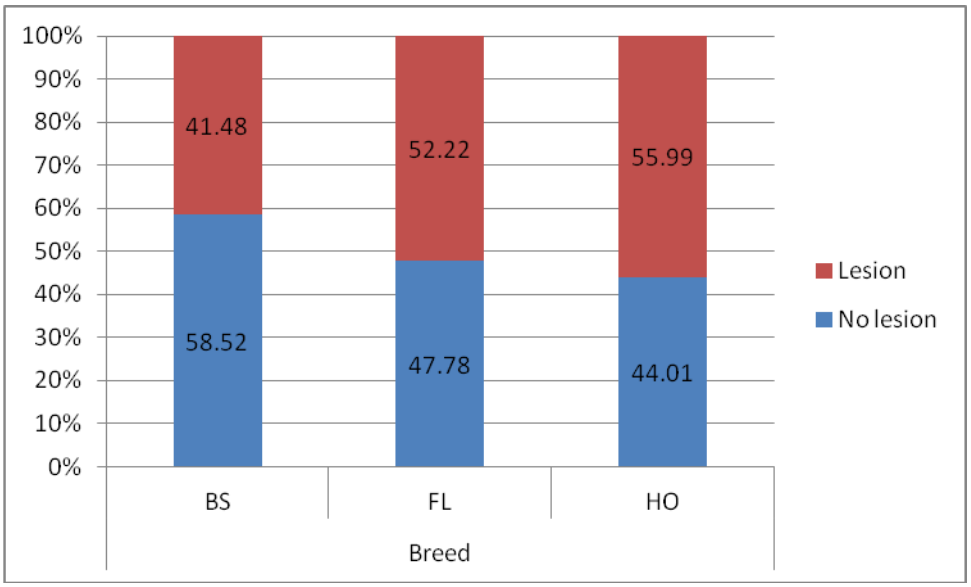

Figure 17: Percentage of records with hoof or leg lesions in function of cow breed.

2.1.3. Parity

Parity has a very strong association with the locomotion scores given to cows. Heifers are sound in 89.73% of cases, while this percentage is reduced to only 56.22% for cows in their 7th parity or more. This is in accordance with findings by DIPPEL *et al.* (2009) who also found that higher parity was associated to higher locomotion scores. In another study (DIPPEL *et al.*, 2009a), they suggest that this is due to the accumulation of risk factors for lameness over the lifetime of the cow: a cow that was once lame has a higher risk of becoming lame again. The Pearson correlation coefficient between parity and locomotion score is 0.238 for this data set. Figure 18 illustrates the association between parity and locomotion scores.

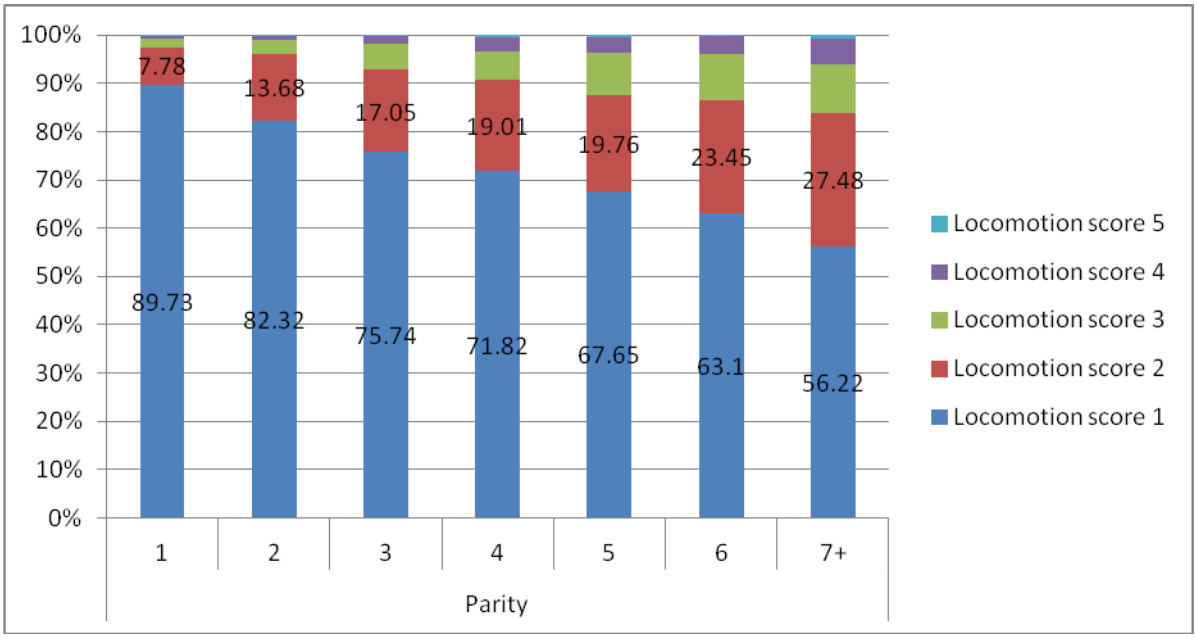

Figure 18: Percentage of records with a certain locomotion score in function of cow parity.

2.1.4. Housing

The 'Efficient Cow' project provided a lot of information about the type of housing the cows were in. Out of this, it was possible to analyse a few influences of housing systems on cow lameness. First of all, cows can be kept in tie-stall or free-stalls. In tied systems, the cows are tethered for certain periods of time. Free-stall systems seem to be slightly better for the cow because of the greater freedom of movement. Indeed, 78% of the cows is sound in free-stall systems against 76% in tie-stalls. However, the percentage of sound cows in tie-stall systems greatly depends on providing exercise to the cow, either in a pasture or a paddock, as shown by Figure 19.

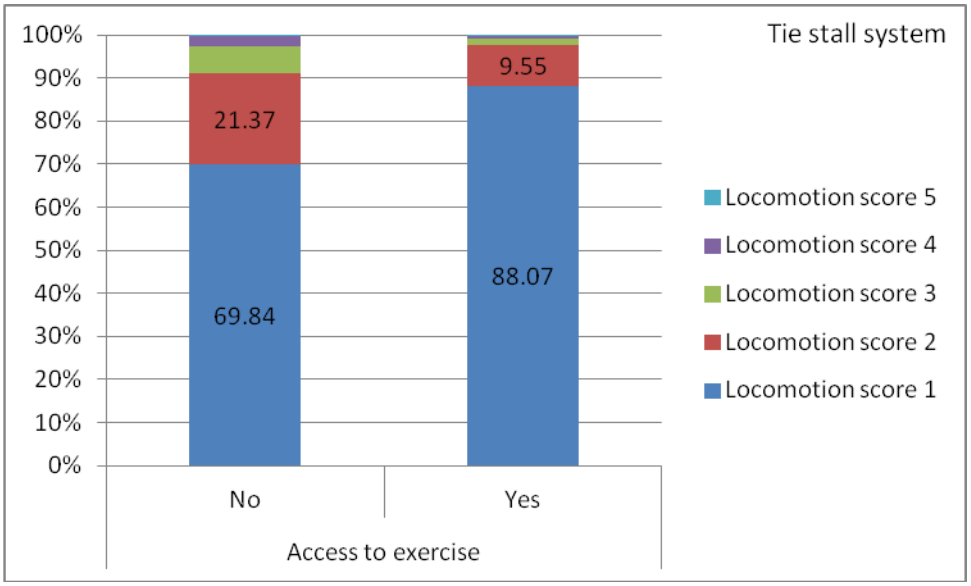

Figure 19: Percentage of records with a certain locomotion score for cows housed in tie-stalls with and without access to exercise (paddock or pasture).

In free-stall systems, exercise has a smaller impact as the cows can move more freely. Still, access to pasture or paddock proved beneficial as shown in Figure 20. The p-value of the chi-square test confirmed an association between access to pasture or paddock and locomotion scores in both tie- and free-stall systems.

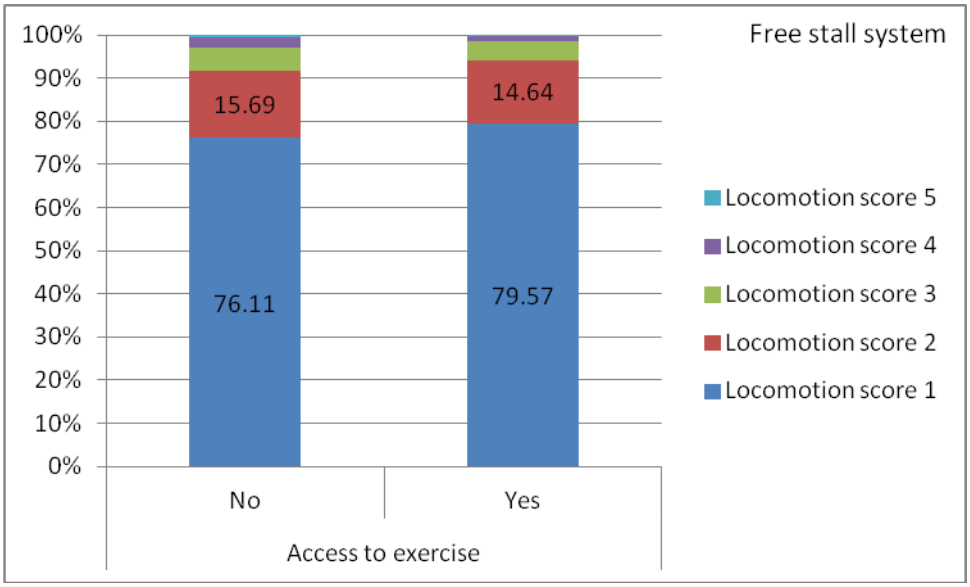

Figure 20: Perecentage of records with a certain locomotion score for cows housed in free-stalls with and without access to exercise (paddock or pasture).

Lameness in free-stall systems was also influenced by the type of free-stall wherein the cows had to rest. For the first test, the records were combined into two distinct groups, cubicle-housed cows, and large deep litter beds. For greater clarity, records from farms with a mix of both systems were not taken into account.

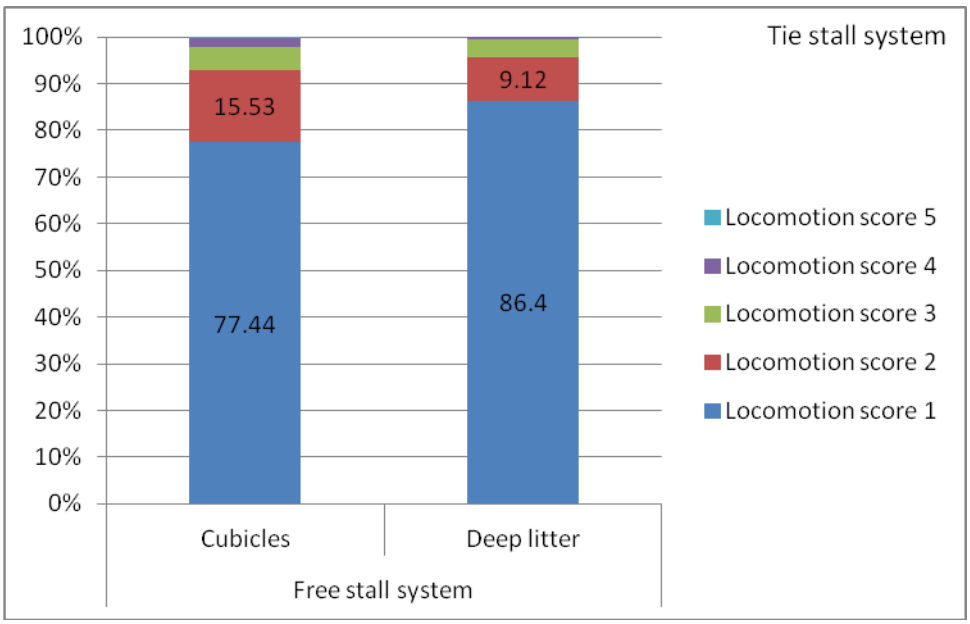

Figure 21: Percentage of records with a certain locomotion score for cows housed in free-stalls with cubicles or with deep litter beds.

It is important to mention that in the present data set, there is great disparity in the number of animals housed in one system or another: 647 records came from deep litter housing against 35458 from cubicle housing, therefore one should not overinterpret the results. However, we found rather large differences between both systems, Figure 21 shows a

difference of almost 10% in sound animals, between cubicle housed cows and deep litter housed cows. The greater freedom of the animals in large litter systems has different advantages for the cow. First of all, it gives her full liberty of movement when lying down or standing up. Secondly, it reduces the amount of objects she could hurt herself against, when evading other cows as shown by the fact that 96% of records with a fracture, dislocation or other leg injury were housed in cubicle systems. Finally, the deep litter causes a dryer floor surface because of a better absorption of urine and manure. Drier and cleaner feet have less risk of developing hoof lesions. An example is dermatitis digitalis which was found in 5.04% of the records housed in cubicle systems against only 0.80% of records in deep litter housings.

2.1.5. Culling reason

Table 13 gives the culling reasons for the 10 590 records that were culled from the herds in the period of data sampling. Apart from claw and leg problems, the only other culling reasons with a percentage of sound animals lower than 70% were old age, mastitis and death. Old age is linked to parity and 3.1.1. showed the association between parity and locomotion score. In the chapter about consequences of lameness, it was already mentioned that lameness can increase the risk for mastitis, which could be an explanation for this slightly lower percentage of sound cows. Interestingly, metabolic diseases do not seem to have poor locomotion scoring. However, this may be due to the way culling reasons are recorded. Indeed, lameness problems because of laminitis can be more visual than the ruminal acidosis that caused it in the first place, which could lead to the culling reason being recorded as lameness instead of a metabolic problem.

Table 13: The locomotion scores given to records in function of their culling reason.

Culling reason	Locomotion score (% of animals)				
	1	2	3	4	5
Mastitis	65.35	22.18	8.39	3.60	0.49
Low yield	73.20	17.69	6.39	2.45	0.27
Old age	52.06	28.49	12.01	6.75	0.69
Infectious diseases	75.76	19.70	3.03	1.52	0.00
Claw and leg problems	42.54	25.19	19.79	10.90	1.59
Slaughter	72.47	20.04	5.29	2.20	0.00
Poor milkability	75.87	16.43	5.59	1.40	0.70
Other reasons	76.07	16.95	5.37	1.43	0.18
Metabolic diseases	73.26	17.48	5.14	3.60	0.51
Infertility	72.98	18.23	5.82	2.60	0.37
Death	45.83	41.67	12.50	0.00	0.00
Sale for breeding	90.54	6.97	1.91	0.50	0.08

2.2. Milk components potentially linked to lameness

2.2.1 Looking at specific animals

The relationship between milk components and lameness, is potentially very complicated. First, milk can change before animals become lame, but also after they are lame. Even if specific cases do not allow drawing definitive conclusions, they help to establish patterns. Therefore, the lactations of two cows were studied more closely. The resulting tables and graphs make it possible to observe the simultaneous evolution of lameness, and the value of biomarkers. Based on these, it is possible to make some suggestions of possible reasons for encountered changes in milk composition.

Cow 1 was locomotion scored 5 times during the first 6 months of her lactation. Figure 22 shows the locomotion scores given over the course of her lactation and the changes in biomarker concentrations. The different biomarkers visibly decrease and increase simultaneously with the locomotion scores. First this example shows that different elements influencing milk composition may lead to confusions. This cow showed high values for BHB, acetone and citrate during the first month associated to her use of a lot of energy leading to body fat mobilization. Similarly, high values were observed much later in her lactation (month 5), a moment where cows normally are not in negative energy balance. Therefore, one could draw hypotheses about the cause-consequence relationship of lameness and ketosis. On the one hand, the increase in lameness biomarker concentration during the 5th month could be the cause of the high locomotion score, although metabolic problems usually take some time to provoke clinical lameness, unless they cause an acute problem, like acute laminitis (MASON, 2008). On the other hand, the higher concentration of body fat mobilization indicators could be the result of an almost total immobilization of the cow with locomotion score 5, which indicates severe lameness (SPRECHER *et al.,* 1997), now unable to feed herself, because of a lesion or earlier metabolic disorder. This cow, however, did not present any hoof lesions during hoof trimming. In which case, the lameness would be due to a metabolic disorder like the one indicated here during the first month.

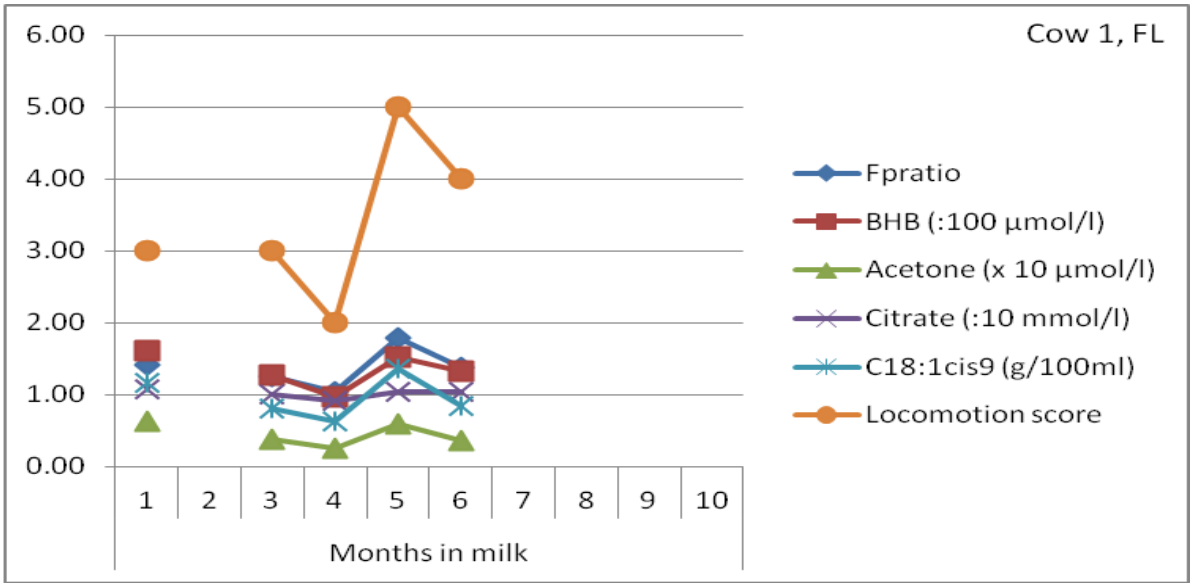

Figure 22: Cow 1; changes in locomotion score and milk composition over the course of the lactation.

Cow 2 (Figure 23) seems to show other aspects of the relationship between lameness and ketosis). From month 4 to month 5, her locomotion score doubles from 2 to 4, while all the body fat mobilization and ketosis indicators drop by about 20 to 40%. So, it could be argued that the lameness in month 5 was caused by the mobilization of body reserves several months before. It could also be that this same mobilization in month 4 had an effect on hoof quality which caused the sole ulcer the cow was treated for in month 6 of her lactation. In month 7, there is again a simultaneous increase of locomotion score and energy body fat mobilization indicators, because a locomotion score of 5 indicates an almost complete immobilization of the cow which can reduce dramatically her feed intake, and cause her body to use energy reserves.

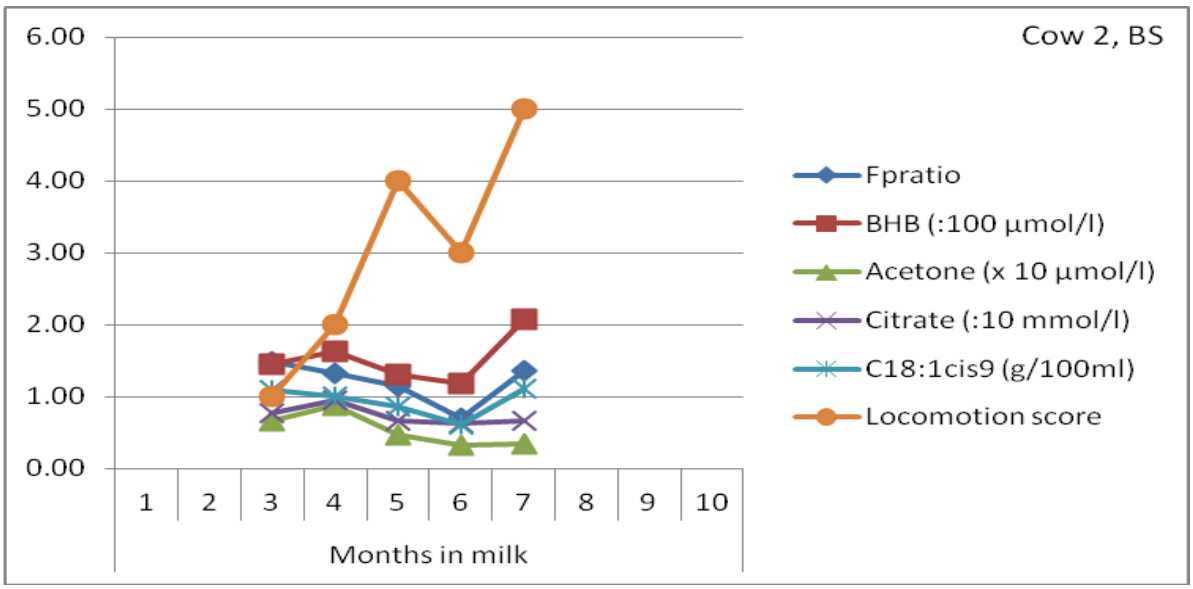

Figure 23: Cow 2; changes in locomotion score and milk composition over the course of the lactation.

So, these are reflections we can derive from looking at ketone bodies, citrate, a fatty acid and fat protein content ratios. Another group of milk components, minerals, have also been linked to hoof quality, and therefore lameness (TOMLINSON *et al.*, 2004; HAMANN & KRÖMKER, 1997), but it was harder to draw conclusions from looking at their changes in concentration, as they do not change as dramatically as that of the previous molecules, as seen in Figure 24 and 25.

HAMANN & KRÖMKER (1997) found a decrease in potassium (K), like seen during the 5th month of cow 1, to be an indicator of acidosis. The ketosis indicators, however, indicated ketosis. These 2 findings are in opposition with each other as acidosis results from an excess in energy intake while ketosis results from the opposite. In this case however, given the very consistent ketosis indicators it is more likely the cow had ketosis than acidosis.

An increase in sodium (Na) or phosphorus (P) concentration can be a sign of Ca deficiency. This increase should be in the range of 70% which was the case for neither of the cows here.

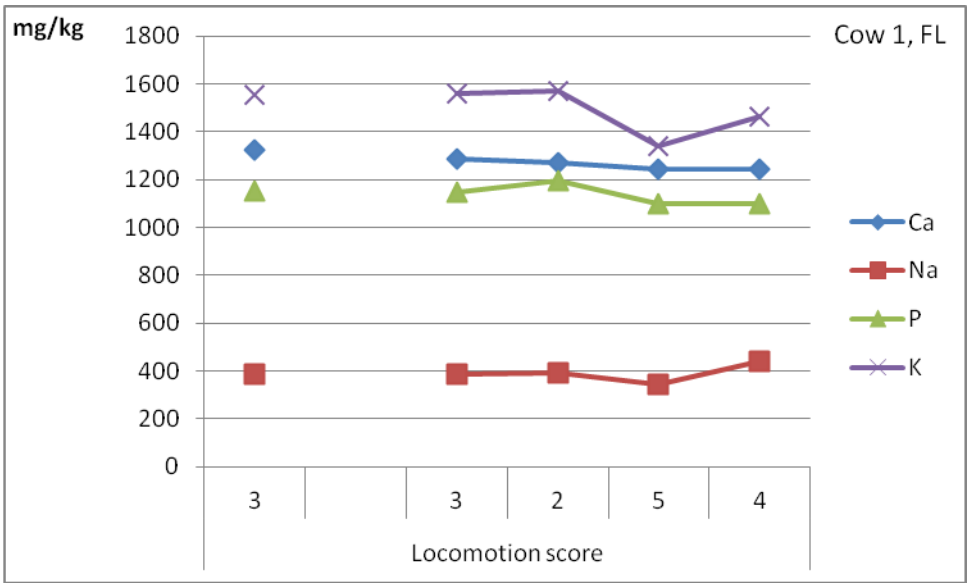

Figure 24: Cow 1; changes in milk composition in function of the successive locomotion scores.

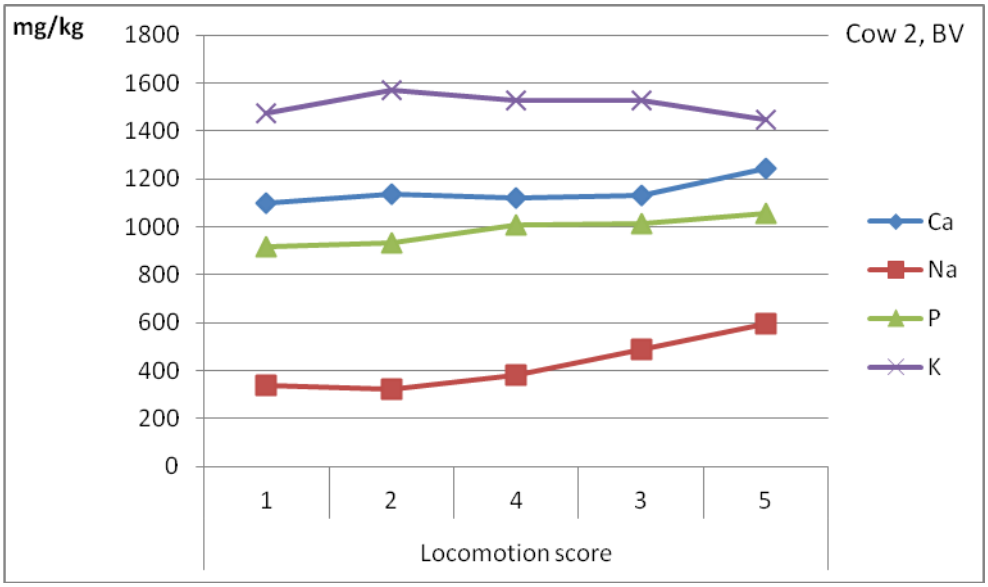

Figure 25: Cow 2; changes in milk composition in function of the successive locomotion scores.

So, these are all assumptions made on the basis of the information obtained from the present data, and the knowledge acquired through the literature review. It again illustrates the complexity of the aetiology of lameness. The relationships between the indicators and condition that is lameness, can vary quite strongly. It would be interesting to have more regular observations to go further into the reflection around this topic. Blood samples could also be of help to confirm certain assumptions made here.

Even if the presented findings are only based on 2 cows, they suggest that there is indeed a relationship between lameness and biomarkers like F:P ratio, BHB, acetone, citrate, fatty acids and minerals. The next title, 2.2.2., examines if these potential relationships between lameness and biomarkers can be generalized to the whole sample.

2.2.2. Generalization

Table 14 shows the difference in biomarker composition between cows with locomotion score 1 and 5. These values were calculated for the whole sample, and are therefore across breeds. Even then, only 37 of the 9324 records had a locomotion score of 5 which explains the higher standard deviation values found for that locomotion score.

Table 14: Mean and standard deviation values of biomarkers for cows of locomotion score 1 and 5.

		Locomotion score 1		Locomotion score 5	
	Unit	Mean	St d	Mean	St d
F:P ratio		1.28	0.21	1.36	0.33
BHB	µmol/l	167.26	49.20	163.59	55.12
Acetone	µmol/l	0.06	0.03	0.06	0.04
Citrate	mmol/l	9.45	1.58	9.18	2.21
C18:1cis9	g/100ml	0.82	0.21	0.93	0.36
LCFA	g/100ml	1.68	0.39	1.86	0.63
MCFA	g/100ml	2.24	0.45	2.23	0.56
SCFA	g/100ml	0.37	0.07	0.38	0.08
Ca	mg/kg	1202.42	98.01	1187.48	107.98
Na	mg/kg	352.78	57.28	1021.96	86.12
P	mg/kg	1028.41	90.44	1497.91	109.55
K	mg/kg	1534.64	92.45	361.54	55.61

Still, some information can be found in this comparison. Interestingly, for this data set the ketone body concentrations (BHB, acetone and citrate) go down instead of up when lameness gets worse. Even when looking at the complete set of locomotion scores (appendix 3), no clear increase is visible. The values first go down for the scores 2 and 3, and then seem to increase again for the scores 4 and 5, but not enough to surpass the values of a sound cow (locomotion score 1). This is in contradiction with what would be expected since body fat mobilization and ketosis are linked to lameness. As discussed during the literature review, several authors (HAMANN & KRÖMKER, 1997; ENJALBERT *et al.*, 2001; MCART *et al.*, 2012; MCART *et al.*, 2013; VAN DER DRIFT *et al.*, 2012; GRELET *et al.*, 2016) found increases in ketone bodies and associated citrate to be good indicators of fat mobilization and ketosis. However, it is important to remember that the effect of metabolic problems on lameness, could be delayed as discussed in 2.2.1. and therefore, high locomotion scores do not always appear at the same time as the metabolic problem that caused them.

The correlations of the biomarkers with locomotion score were also computed. Because of the results in the above table, it is not surprising to find a negative correlation between the ketosis indicators and lameness (Table 15), even though generally speaking, a positive correlation would be expected.

Table 15: Correlations between ketosis indicators and locomotion score.
The correlation coefficients are above the diagonal and the corresponding p-values can be found under it.

		Pearson correlation coefficient			
		1	2	3	4
P-value	1. Locomotion score		-0.04923	-0.00331	-0.08204
	2. BHB	<.0001		0.71193	0.5443
	3. Acetone	0.749	<.0001		0.49081
	4. Citrate	<.0001	<.0001	<.0001	

In Table 16, the concentration in medium chain fatty acids (MCFA) seems to decrease a little in favour of a big increase in long chain fatty acids (LCFA) among which C18:1cis9 which corresponds to the findings of VAN HAELST *et al.* (2008), while short chain fatty acids (SCFA) fluctuate a little but stay fairly stable across locomotion scores. The fat to protein ratio also increases as expected (KOFLER *et al.*, 2013; HAMANN & KRÖMKER, 1997; LOKER *et al.*, 2012). The correlations that were found for these biomarkers reflect these relations.

Table 16: Correlations between body fat mobilization indicators and locomotion score.
The correlation coefficients are above the diagonal and the corresponding p-values can be found under it.

		Pearson correlation coefficient					
		1	2	3	4	5	6
P-value	1. Locomotion score		0.04429	0.04617	0.03958	-0.01783	-0.01054
	2. F:P ratio	<.0001		0.67593	0.74404	0.58131	0.6516
	3. C18:1cis9	<.0001	<.0001		0.96831	0.24167	0.2574
	4. LCFA	0.0001	<.0001	<.0001		0.35292	0.42299
	5. MCFA	0.0851	<.0001	<.0001	<.0001		0.89907
	6. SCFA	0.3088	<.0001	<.0001	<.0001	<.0001	

Calcium (Ca) decreases a little, but far less dramatically than sodium (Na), or phosphorus (P) increase and potassium (K) decreases. It is logical that Na and P increase when Ca decreases, as they are indicators of hypocalcaemia (HAMANN & KRÖMKER, 1997).

Table 17: Correlations between minerals and locomotion score.
The correlation coefficients are above the diagonal and the corresponding p-values can be found under it.

		Pearson correlation coefficient				
		1	2	3	4	5
P-value	1. Locomotion score		-0.05665	-0.05432	-0.07626	0.05565
	2. Ca	<.0001		0.47656	-0.32696	-0.00788
	3. P	<.0001	<.0001		0.19516	-0.04616
	4. K	<.0001	<.0001	<.0001		-0.15234
	5. Na	<.0001	0.4467	<.0001	<.0001	

On the whole, the correlations for all biomarkers with lameness are weak. This could be the result of the indirect way in which the feet health is connected to milk composition in comparison to utter health. One could hypothesise that, to have a repercussion on milk composition, lameness first has to influence the blood sufficiently, or be caused by a disorder that influences the blood enough. So that these changes in blood composition may be

'translated' into milk composition. A bacterial infection in the utter, on the other hand, directly influences milk composition. So, this indirect way of connecting milk composition and lameness, could be an explanation of the low correlations.

Nevertheless, points 2.2.1. and 2.2.1. still seem to point towards, if not a very clear relationship, at least some relationship between the molecules chosen as biomarkers in this work, and lameness. This supports the hypothesis that there might be benefits in adding these traits in the oriented MIR calibration of the second study.

3. First study: Classic MIR calibration

3.1. Animal Science Day article (adapted from MINEUR *et al.*, 2017)

In Table 18, we can see the results for different data sets. The sensitivity, i.e. lame animals predicted as lame by the model, and the specificity, i.e. non-lame animals predicted as non-lame, is presented for each data set for calibration and validation (PENN STATE EBERLY COLLEGE OF SCIENCE, 2017). Results are presented by subsets. The complete data set ('All') had the lowest RMSEcv for 11 latent variables. Because all other data sets were then compared to the data set 'All', the same number of 11 latent variables was chosen for every set. For each data set, the first row is the results obtained when using the first derivative, the second row when using the second derivative.

Table 18: Number of records (N), number of lame records (Lame), sensitivity and specificity for calibration, validation and for first derivative (first row) or second derivative (second row) for different data sets.
*HHE = Heelhorn erosion, **WL = white line defect.

Subset	N	Lame	Calibration		Validation	
			Sensitivity (%)	Specificity (%)	Sensitivity (%)	Specificity (%)
All	9811	795	63	63	60	62
			61	65	54	65
First half of lactation	5509	490	60	68	53	66
			59	69	52	67
Last half of lactation	4302	305	68	64	60	64
			61	68	56	67
First third of lactation	3806	348	70	55	67	58
			68	60	62	61
Last third of lactation	2479	176	65	69	45	66
			57	73	32	68
Simmental	6828	578	71	62	62	59
			68	66	55	61
Holstein	1560	121	68	70	43	71
			55	77	43	79
Brown Swiss	1423	96	68	70	67	63
			70	73	59	68
Heifer (parity = 1)	2792	96	73	67	56	65
			64	72	47	70
Young (parity = 1 or 2)	4855	195	71	59	49	58
			64	63	47	62
Old (parity > 2)	4956	600	68	60	60	61
			68	62	60	62
HHE*, 3 weeks	596	52	88	93	85	91
			84	93	80	92
HHE* & lame, 3 weeks	273	52	87	92	85	86
			87	91	85	88
WL**, 3 weeks	678	41	59	91	42	88
			45	93	25	91
WL** & lame, 3 weeks	465	41	81	89	53	84
			81	89	53	84

3.1.1. Spectral data pre-treatment

As we can see in Table 18, changing the pre-processing from a first derivative and SNV to a second derivative and SNV produces varying results, depending on the subsets. However, using the second derivative does seem to have the tendency to favour results for true negatives, i.e. truly sound cows, while obtaining lesser results for true positives, i.e. truly lame cows. However, practical usefulness of true positives is much more important than true negatives, because it is more important for the farmer to know which animals have a problem and need treatment, than knowing which animals do not. For this reason, we will focus on the first derivative in the results reported in Table 18 and in the following studies, titles 3.2. and 4., no second derivatives were used.

3.1.2. Calibration results

Results are split between calibration and validation results (Table 18). For the reason explained earlier, only results for the first derivative are discussed. The results for the complete number of records, 'All', were not very high. In the calibration context, sensitivities and specificities were 63%; validation results 60% and 62%. This could indicate that many other factors were influencing lameness. Results were obtained for different lactation stages, breeds, and lactation numbers. We will focus on validation results.

Separating the records into 2 halves of lactation, i.e. first 5 months and 5 last months, gives mitigated results with a little improvement for the last 5 months, and worse results for the first 5. The first 5 months of the lactation of a cow comprise two very different periods. The first 2 months are characterized by great energy expenditure, and possible negative energy balance as the cow has just calved, and her production increases towards a peak at approximately 2 months (60 days). After this, her milk yield gradually decreases and, within the next months, she should build her energy reserves up again as she reaches her positive energy balance. The results obtained in the first half of lactation, might therefore be a repercussion of mixing those two very different metabolic stages in the cow lactation.

When we look now at the first 100 days (First third of lactation), the results are clearly better. This supports the hypothesis that the link between MIR spectra and lameness, is lactation stage dependent. The bad results for the last third of lactation might be linked to the fact that there can be great variation in the last hundred days, with some cows obliged to finish their lactations earlier because of health reasons or otherwise. In conclusion, it might therefore be better to establish lactation stage specific prediction equations. Recently, this type of strategy was used successfully for MIR predicted methane emission.

Separating breeds gave mixed results. It improved for Simmental and Brown Swiss, but deteriorated the results for Holstein cows. There are no obvious reasons why Holstein did not perform as well, but these results showed that the link between lameness and spectra

seems to be breed specific. Calibration equation should be developed for each breed separately.

In point 2.1.3., it was seen that parity was greatly associated with the development of lameness and this is reflected in the calibration results as well. All three subsets, i.e. heifers, young cows and old cows, obtained better results for calibration. The biggest improvement in calibration was obtained for the heifers. Surprisingly, however, the results deteriorated for validation. Validation also deteriorated for young cows, and did not really improve for old cows. A possible explanation for this difference in improvement could be the disparity between the numbers of non-lame compared to lame record exacerbated by the smaller size of the validation set. When looking at calibration results again, it seems like isolating a parity on its own (heifers), delivers better results than combining even only 2 parities together (young). Again, these findings show the MIR based equations might be developed for animals in a given parity.

Lameness can have many origins, however creating subsets based on a specific disease, narrows down the cause of lameness and its effect on milk composition. Results (Table 18) were good especially for heel horn erosion. This data set is also one of the sets with the highest proportion of lame animals which could also explain the better results. The white line disorder set on the other hand, shows the risk of using a too small dataset of which not enough records are scored lame. It gives better results again when the records used are of only lame and suffering from a white line disorder, thus eliminating the records suffering from the lesion, but not displaying any lameness which could confuse the model.

3.2. Additional computations

The complete data set 'Detection' had the lowest RMSEcv for 11 latent variables. So, a model with 11 latent variables was chosen for all subsets as well to make comparison easier.

Table 19: Sensitivity and specificity for calibration, validation and for first derivative (first row) or second derivative (second row) for different 'Detection' subsets.

DETECTION	Calibration		Validation	
Subset	Sensitivity (%)	Specificity (%)	Sensitivity (%)	Specificity (%)
Detection	84	67	47	75
First 2 months of lactation	82	75	17	72
Last 8 months of lactation	62	77	66	65
Simmental	62	77	44	76
Brown Swiss	100	93	NaN	80
Holstein	49	81	0	87
Heifer (parity = 1)	100	96	0	81
Young (parity = 1 or 2)	74	83	56	73
Old (parity > 2)	77	87	42	76

At first glance, the results for the 'Detection' subsets, computed in Table 9, are totally disparate. Some subsets like Heifer have a perfect result for calibration sensitivity, but utterly fail to predict any lame animal correctly during validation. Looking at the descriptive tables in material and methods, it seems like, in this case, there were many more records that fell into the 'Prediction' category than the 'Detection' category. Separating per farm also seems to have a bigger effect on smaller sets. As the number of lame animals per farm can vary, this could potentially influence the balance between lame and non-lame records more strongly than randomly selecting per record. Very small data sets like the Brown Swiss or Holstein are not adequate for prediction. The NaN in the validation set of the Brown Swiss data is caused by an absence of lame animals in that subset.

As the complete 'Prediction' set obtained the best results for a model with 15 latent variables, the same number of latent variables was chosen for all the subsets.

Table 20: Sensitivity and specificity for calibration, validation and for first derivative (first row) or second derivative (second row) for different 'Prediction' subsets.

PREDICTION	Calibration		Validation	
Subset	Sensitivity (%)	Specificity (%)	Sensitivity (%)	Specificity (%)
Prediction	72	62	56	62
First 2 months of lactation	67	68	38	68
Last 8 months of lactation	74	62	63	58
Simmental	69	67	60	64
Brown Swiss	71	73	60	65
Holstein	73	82	33	70
Heifer (parity = 1)	67	78	24	74
Young (parity = 1 or 2)	68	65	53	59
Old (parity > 2)	61	67	53	71
Locomotion score 1, 3, 4 and 5	64	69	62	67
Locomotion score 1, 4 and 5	75	72	62	66
Balanced locomotion score 1	64	63	42	70

'Prediction' subsets contained more records, and were more balanced in lame and non-lame animals than their counterparts in 'Detection', which makes a more precise interpretation of the results possible. Separating the duration of lactation in function of the lactation peak at two months, was definitely an improvement for the last 8 months. The results also improved for the calibration of the first 2 months, but not for the validation. This could again be due to a smaller proportion of lame records in that set, or to the great variability between the records that compose it.

The results obtained for separating the breeds were a little lower for calibration, and slightly higher for validation than the full set, but fairly similar to the previous study. Although in this case, the Holstein records fared better, and the Simmental ones a little less than previously.

Separating the parities did not have clearly beneficial effects on the results, mostly improving the specificity, but not the sensitivity. Sensitivity, however, is the most interesting value to have high results for, as that value indicates when animals are lame, and need treatment.

Balancing out the data by reducing the number of locomotion scores 1 did not result in a great improvement. On the other hand, deleting score 1, or the scores 1 and 2, seemed to have a beneficial effect on both sensitivity and specificity. This is probably because some of the 'in-between' records that can cause confusion in the model were eliminated, e.g. a cow that is not really lame, but got a 2 because she was stiff from lying down for a long time. So, it seems that clearer stricter definitions of lameness, make it easier for the model.

This suggests that the complexity of the link between lameness and milk composition, is probably what makes the prediction model stagnate around 60 to 70% at best, for the results in sensitivity and specificity.

A general remark is that separating 'Detection' and 'Prediction' probably does not give the improvement hoped for. This is due to the complexity of the relationships between the causes of lameness, its impacts and the milk composition. This complexity is probably there, regardless whether the MIR sample was taken before or after locomotion scoring.

4. Second study: Oriented MIR calibration

Because of the full 'Prediction' set, 15 latent variables were chosen for all subsets. The improvement of the results, reported in Table 21 seems to stop again around 60 to 70% for calibration and 50 to 60% for validation. The results were better for fatty acids than for the ketosis molecules. This reflects the better correlations found, as well as at the beginning of this chapter. It had a neutral effect. Combining biomarkers seemed to improve the calibration even a little more, but did not improve the validation.

Table 21: Results in sensitivity and specificity for calibration and validation of oriented MIR calibration for various subsets.

PREDICTION	Calibration		Validation	
Subset	Sensitivity (%)	Specificity (%)	Sensitivity (%)	Specificity (%)
spectra	68	66	53	64
spectra + BHB	67	65	53	64
spectra + acetone	67	65	14	93
spectra + citrate	69	66	52	63
spectra + BHB, acetone, citrate	69	65	51	64
spectra + C18:1cis9	70	65	52	64
spectra + LCFA	70	65	52	64
spectra + MCFA	73	65	54	64
spectra + SCFA	69	66	54	64
spectra + MCFA, SCFA	70	65	53	64
spectra + C18:1cis9, LCFA, MCFA, SCFA	69	65	51	64
spectra + BHB, acetone, citrate, C18:1cis9, LCFA, MCFA, SCFA	68	66	53	65
spectra + Ca	69	66	51	63
spectra + all biomarkers	67	66	50	64

CHAPTER V: CONCLUSIONS AND PERSPECTIVES

The aim of this work was to test the feasibility of a MIR based prediction equation. To this end, PLS-DA calibrations were applied to MIR spectral data. Results showed that the use of milk MIR spectra with the aim of detecting lameness in cows, still needs additional research.

Increasing the number of true positives (i.e. sensitivity) at validation is rather critical, as for the farmer it is more important to detect lame animals in order to treat them. However, in most cases, this percentage is still too low: not predicting 30 to 40% of the cows that need help, is very inefficient. Furthermore, by chance, we have a 50% chance of classifying an animal correctly in the lame or non-lame categories, as there are only two options; predicting 60 or even 70% correctly due to the model seems to be only a small improvement compared to chance. Increasing the number of true positives is also essential as it would not be useful to spend time checking cows for lameness problems they do not have. There are several ways that could help achieve these results.

First, models need to define precisely the target to be predicted, and be more refined to take into account the different sources of variation that exist in the field, as only the most homogeneous datasets, about heel horn erosion and white line disorder, produced results starting to be interesting. This suggests that potentially the MIR technology has to be used for very specific situations, and that not all types of lameness can be predicted. Moreover, the different sources of variation need to be better controlled before the technology can be used on a larger scale, with data coming from varied animals and farms.

It could also be of interest to look more closely, and with smaller intervals between measures at a group of cows of varied natures, i.e. different breeds, parities, lactation durations. In the present data set, there was a lot of information about a large number of animals, but at the cow level, there was only an average of 2.4 records per cow. To better understand the causes of lameness, it could be very valuable to observe the smaller changes in value, during shorter periods of time of the biomarkers discussed in this work.

Then, there is the possibility to add supplementary information coming from new variables like BCS, milk yield or diet for each record. During the course of this work, the mobilization of body fat reserves and lameness have often been linked to each other. Seeing as BCS is a visual assessment of a cows body condition, and therefore fat reserves, adding it to the model could provide more information about the cows physiological condition, especially if it is recorded on the same day as the test day milk, used afterwards for MIR analysis. Additionally, different breeds can have different average body weights, in which case, BCS could increase the calibration results inside breeds. Cows with higher milk yield are more susceptible to lameness, and milk yields also vary between breeds. Finally, the diet was not discussed in detail in the present work, but is was put into relationship with subacute ruminal acidosis and the potentially ensuing laminitis.

In conclusion, there are good reasons to develop a detection of lameness equation. Determining early lameness accurately ,and treating the concerned cow could prevent light lameness stages from worsening, and having long-lasting consequences on cow health and profitability. This study shows a complex link between cow lameness and milk composition, but a link nonetheless which could be studied in further works.

REFERENCES

Anon, 2017a. Foot and claw physiology. *Agriculture and Horticulture Development Board, Dairy*. Available at: https://dairy.ahdb.org.uk/technical-information/animal-health-welfare/lameness/physiology/#.WYCVX4TyiM8.

Anon, 2017b. Hoof - Anatomy & Physiology. *WikiVet*. Available at: https://en.wikivet.net/Hoof_-_Anatomy_%26_Physiology.

Anon, 2017c. Projekte Efficient Cow. *Rinderzucht Austria*. Available at: https://www.zar.at/Projekte/Efficient-Cow.html.

Archer, S.C., Green, M.J. & Huxley, J.N., 2011. Association between milk yield and serial locomotion score assessments in UK dairy cows. *Journal of dairy science*, 93(9), pp.4045–4053.

ASHWOOD, A., 2011. Structure and lameness. *BRAHMAN NEWS*. Available at: http://www.brahman.com.au/technical_information/selection/structureAndLameness.html.

Barker, Z.E. et al., 2010. Assessment of lameness prevalence and associated risk factors in dairy herds in England and Wales. *Journal of Dairy Science*, 93(3), pp.932–941.

Berry, S.L., Cook, N.B. & Prairie, E., 2017. Locomotion Scoring of Dairy Cattle. *Availa4*, p.3.

Bicalho, R.C., Machado, V.S. & Caixeta, L.S., 2009. Lameness in dairy cattle: A debilitating disease or a disease of debilitated cattle? A cross-sectional study of lameness prevalence and thickness of the digital cushion. *Journal of dairy science*, 92(7), pp.3175–84.

Braun, R.K. et al., 1986. Body condition scoring dairy cows as a herd management tool. *Compendium on Continuing Education for the Practicing Veterinarian*, 8, pp.F62–F67.

Bruijnis, M.R.N., Hogeveen, H. & Stassen, E.N., 2010. Assessing economic consequences of foot disorders in dairy cattle using a dynamic stochastic simulation model. *Journal of Dairy Science*, 93(6), pp.2419–2432.

Channon, A.J. et al., 2009. Variability of Manson and Leaver locomotion scores assigned to dairy cows by different observers. *The Veterinary Record*, 164(13), pp.388–392.

Chapinal, N. et al., 2013. Short communication: Herd-level reproductive performance and its relationship with lameness and leg injuries in freestall dairy herds in the northeastern United States. *Journal of dairy science*, 96(11), pp.7066–72.

Charfeddine, N. & Pérez-Cabal, M.A., 2017. Effect of claw disorders on milk production, fertility, and longevity, and their economic impact in Spanish Holstein cows. *Journal of Dairy Science*, 100(1), pp.653–665.

Clarkson, M.J. et al., 1996. Incidence and prevalence of lameness in dairy cattle. *The Veterinary Record*, 138(23), pp.563–567.

Collard, B.L. et al., 2000. Relationships between energy balance and health traits of dairy cattle in early lactation. *Journal of dairy science*, 83(11), pp.2683–2690.

Cook, N.B., Nordlund, K. V. & Oetzel, G.R., 2004. Environmental Influences on Claw Horn Lesions Associated with Laminitis and Subacute Ruminal Acidosis in Dairy Cows. *Journal of Dairy Science*, 87, pp.E36–E46.

CRA-W, 2017. Training in Vibrational Spectroscopy and Chemometrics. In Gembloux, p. 1095.

Daly, R., 2014a. Lameness in Cattle: Causes Associated with Injury. *IGrow A Service of SDSU Extension*. Available at: http://igrow.org/livestock/beef/lameness-in-cattle-causes-associated-with-injury/.

Daly, R., 2014b. Treating Lameness in Cattle. *IGrow A Service of SDSU Extension*. Available

at: http://igrow.org/livestock/beef/treating-lameness-in-cattle/.
Dippel, S. et al., 2009a. Risk factors for lameness in cubicle housed Austrian Simmental dairy cows. *Preventive Veterinary Medicine*, 90(1–2), pp.102–112.
Dippel, S. et al., 2009b. Risk factors for lameness in freestall-housed dairy cows across two breeds, farming systems, and countries. *Journal of Dairy Science*, 92, pp.5476–5486.
van der Drift, S.G. a. et al., 2012. Genetic and nongenetic variation in plasma and milk β-hydroxybutyrate and milk acetone concentrations of early-lactation dairy cows. *Journal of Dairy Science*, 95(11), pp.6781–6787.
Egger-Danner, C. et al., 2015. *ICAR Claw Health Atlas* J. Cole & J. Kofler, eds., Rome: ICAR, Italy.
Enjalbert, F. et al., 2001. Ketone bodies in milk and blood of dairy cows: relationship between concentrations and utilization for detection of subclinical ketosis. *Journal of dairy science*, 84(3), pp.583–589.
Espejo, L. a, Endres, M.I. & Salfer, J. a, 2006. Prevalence of lameness in high-producing holstein cows housed in freestall barns in Minnesota. *Journal of dairy science*, 89(8), pp.3052–3058.
European Commission, 2012. *EU dairy farms report 2012, based on FADN data*,
Garry, F., 2017. Laminitis. *Colorado State University - Veterinary Teaching Hospital*, pp.5–7.
Gengler, N. et al., 2016. Capitalizing on fine milk composition for breeding and management of dairy cows. *Journal of Dairy Science*, 99(5), pp.4071–4079.
Gengler, N. & Consortium, G., 2017. Targeted combination of estimated breeding values for lower accuracy mid-infrared biomarkers increases their usefulness in genetic evaluation of dairy cattle. In *ICAR Meeting Edinburgh*.
Goff, J.P., 2008. The monitoring, prevention, and treatment of milk fever and subclinical hypocalcemia in dairy cows. *Veterinary Journal*, 176(1), pp.50–57.
Gooch, C.A., 2003. Flooring Considerations for Dairy Cows. *Natural Resource, Agriculture, and Engineering Service*, 148.
Gozho, G.N., 2005. *Effects of subacute ruminal acidosis (SARA) on ruminal lipopolysacharide release and on inflammation in cattle*. University of Manitoba.
Grelet, C. et al., 2016. Development of Fourier transform mid-infrared calibrations to predict acetone, β-hydroxybutyrate, and citrate contents in bovine milk through a European dairy network. *Journal of Dairy Science*, 99(6), pp.4816–4825.
Grelet, C. et al., 2012. OptiMIR : Standardisation de spectres moyen infrarouge de lait de vache en vue de la création de nouveaux outils de management d ' exploitation. *3R*, (1), p.5030.
Grelet, C. et al., 2015. Standardization of milk mid-infrared spectra from a European dairy network. *Journal of dairy science*, 98(4), pp.2150–60.
Groen, A.F. et al., 1997. 'Genetic Improvement of Functional Traits in cattle' (GIFT). *EU Concerted Action, Annual report*, p.12.
Van Haelst, Y.N.T. et al., 2008. Short Communication: Elevated Concentrations of Oleic Acid and Long-Chain Fatty Acids in Milk Fat of Multiparous Subclinical Ketotic Cows. *Journal of Dairy Science*, 91(12), pp.4683–4686.
Hamann, J. & Krömker, V., 1997. Potential of specific milk composition variables for cow health management. *Livestock Production Science*, 48(3), pp.201–208.
Haskell, M.J. et al., 2006. Housing System , Milk Production , and Zero-Grazing Effects on Lameness and Leg Injury in Dairy Cows. *Journal of Dairy Science*, 89(11), pp.4259–4266.
Herdt, T.H., 2017. Overview of Ketosis in Cattle. *Merck Veterinary Manual*. Available at: http://www.merckvetmanual.com/metabolic-disorders/ketosis-in-cattle/overview-of-ketosis-in-cattle.

Hernandez, J., Shearer, J.K. & Webb, D.W., 2002. Effect of lameness on milk yield in dairy cows. *Journal of the American Veterinary Medical Association*, 220(5), pp.640–644.

Ito, K. et al., 2010. Lying behavior as an indicator of lameness in dairy cows. *Journal of dairy science*, 93(8), pp.3553–3560.

King, M.T.M. et al., 2017. Cow-level associations of lameness, behavior, and milk yield of cows milked in automated systems. *Journal of Dairy Science*, 100(6), pp.4818–4828.

Kofler, J., 2013. „ Was kosten lahme Kühe ?“.

Kofler, J., 2014. Diagnoseschlüssel zu Klauenerkrankungen Lahmheitsbeurteilung (Locomotion Scoring).

Kofler, J., 2015. Seminar des Ausschusses für Genetik der ZAR: Gesunde Klauen und gute Fundamente Einflussfaktoren und Verbesserungsmaßnahmen. In *Klauenerkrankungen in Österreich – wirtschaftliche Aspekte, Häufigkeiten, Erkennung und fütterungsbedingte Ursachen*. Salzburg, p. 26.

Kofler, J., Keplinger, J. & Rinner, D., 2013. Klauengesundheit bei mutterkühen – reheklauen, wanddefekte und hornspalten stehen im vordergrund. *Wiener Tierärztlichen Monatsschrift*, pp.115–126.

Kossaibati, M. a. & Esslemont, R.J., 1997. The costs of production diseases in dairy herds in England. *The Veterinary Journal*, 154(1), pp.41–51.

Larousse, D. de français, 2017. métabolisme. *Larousse, Dictionnaires de français*. Available at: http://www.larousse.fr/dictionnaires/francais/métabolisme/50809.

Loker, S. et al., 2012. Genetic and environmental relationships between body condition score and milk production traits in Canadian Holsteins. *Journal of dairy science*, 95(1), pp.410–9.

De Marchi, M. et al., 2014. Invited review: Mid-infrared spectroscopy as phenotyping tool for milk traits1. *Journal of Dairy Science*, 97(3), pp.1171–1186.

Mason, S., 2008. The three phases of laminitis. *Dairy Cattle Hoof Health*. Available at: http://dairyhoofhealth.info/lesions/three-phases-laminitis/.

McArt, J.A.A. et al., 2013. Elevated non-esterified fatty acids and β-hydroxybutyrate and their association with transition dairy cow performance. *Veterinary Journal*, 198(3), pp.560–570.

McArt, J.A.A., Nydam, D. V & Oetzel, G.R., 2012. Epidemiology of subclinical ketosis in early lactation dairy cattle. *Journal of dairy science*, 95(9), pp.5056–66.

Metzner, M., Heuwieser, W. & Klee, W., 1993. Die Beurteilung der Körperkondition (body condition scoring) im Herdenmanagement. *Der praktische Tierarzt*, 74, pp.991–8.

Newsome, R.F. et al., 2017. A prospective cohort study of digital cushion and corium thickness. Part 1: Associations with body condition, lesion incidence, and proximity to calving. *Journal of Dairy Science*, 100(6), pp.4745–4758.

Oetzel, G.R., 2015. Subacute Ruminal Acidosis in Dairy Herds : Physiology , Pathophysiology , Milk Fat Responses , and Nutritional Management. *University of Wisconsin, School of Veterinary Medicine*, pp.89–119.

Plaizier, J.C. et al., 2008. Subacute ruminal acidosis in dairy cows: The physiological causes, incidence and consequences. *Veterinary Journal*, 176(1), pp.21–31.

Räber, M. et al., 2004. The bovine digital cushion - A descriptive anatomical study. *Veterinary Journal*, 167(3), pp.258–264.

RINDERZUCHT AUSTRIA, 2017a. Brown Swiss. Available at: http://en.zar.at/Cattle_breeding_in_Austria/Cattle_breeds/Brown_Swiss.html.

RINDERZUCHT AUSTRIA, 2017b. Efficient Cow, Die richtige Kuh für meinen Betrieb. , p.19.

RINDERZUCHT AUSTRIA, 2017c. Holstein. Available at: http://en.zar.at/Cattle_breeding_in_Austria/Cattle_breeds/Holstein.html.

RINDERZUCHT AUSTRIA, 2017d. Simmentals. Available at: http://en.zar.at/Cattle_breeding_in_Austria/Cattle_breeds/Fleckvieh.html.

Rinnan, Å. et al., 2009. *Data Pre-Processing* 1st ed., Elsevier Inc.

Rodriguez, L. & DeFrain, J., 2017. Understanding the Anatomy of Cows Hooves.

Rollin, B.E., 2005. An ethicist's commentary on trimming of lame dairy cattle. *Veterinary medical ethics déontologie vétérinaire*, (March), pp.483–484.

Rouha-Mülleder, C. et al., 2009. Relative importance of factors influencing the prevalence of lameness in Austrian cubicle loose-housed dairy cows. *Preventive Veterinary Medicine*, 92(1–2), pp.123–133.

Rutherford, K.M.D. et al., 2009. Lameness prevalence and risk factors in organic and non-organic dairy herds in the United Kingdom. *The Veterinary Journal*, 180(1), pp.95–105.

SAS Institute Inc., 2017. SAS 9.4.

Smith, B.C., 1996. *Fundamentals of Fourier Transform Infrared Spectroscopy* 2nd editio., New-York: CRC Press.

Sogstad, Å.M., Østerås, O. & Fjeldaas, T., 2006. Bovine Claw and Limb Disorders Related to Reproductive Performance and Production Diseases. *Journal of Dairy Science*, 89(7), pp.2519–2528.

Soyeurt, H. et al., 2011. Mid-infrared prediction of bovine milk fatty acids across multiple breeds, production systems, and countries. *Journal of Dairy Science*, 94(4), pp.1657–1667.

Sprecher, D.J., Hostetler, D.E. & Kaneene, J.B., 1997. A lameness scoring system that uses posture and gait to predict dairy cattle reproductive performance. *Theriogenology*, (47), pp.1179–1187.

The MathWorks Inc., 2000. MATLAB 6.1.

van der Tol, P.P.J. et al., 2004. The effect of preventive trimming on weight bearing and force balance on the claws of dairy cattle. *Journal of Dairy Science*, 87(6), pp.1732–1738.

van der Tol, P.P.J. et al., 2002. The pressure distribution under the bovine claw during square standing on a flat substrate. *Journal of Dairy Science*, 85(6), pp.1476–1481.

Tomlinson, D.J., Mülling, C.H. & Fakler, T.M., 2004. Invited Review : Formation of Keratins in the Bovine Claw : Roles of Hormones , Minerals , and Vitamins in Functional Claw Integrity. , pp.797–809.

Tomlinson, D.J., Mülling, C.H. & Fakler, T.M., 2004. Invited Review: Formation of keratins in the bovine claw: roles of hormones, minerals, and vitamins in functional claw integrity. *Journal of Dairy Science*, 87(4), pp.797–809.

Vermunt, J.J., 1992. 'Subclinical' laminitis in dairy cattle. *New Zealand Veterinary Journal*, 40(4), pp.133–138.

Whay, H.R. et al., 1998. The influence of lesion type on the duration of hyperalgesia associated with hindlimb lameness in dairy cattle. *Veterinary Journal*, 156(1), pp.23–29.

Whay, H.R., Waterman, A.E. & Webster, A.J.F., 1997. Associations between locomotion, claw lesions andnociceptive threshold in dairy heifers during the peri-partum period. *The Veterinary Journal*, 154(2), pp.155–161.

Zottl, K. et al., 2017. On-farm recording of novel traits in dairy cattle - genetic parameters and recommendations. In S. Niklitschek et al., eds. *ICAR Technical Series: A global perspective on performance recording and animal identification*. Rome: ICAR, pp. 183–188.

Table 21: Prevalence of lesions found in different studies by KOFLER *et al.* (2013), WHAY *et al.* (1998), and 4 studies cited by COOK *et al.* (2004). Those 4 studies were done by MURRAY *et al.*, KOSSAIBATI AND ESSLEMONT, WARNICK *et al.* and COOK *et al.* Hol stands for Holstein Friesian. The unmentioned breed in the studies in the UK and USA are quite likely to be Holstein Friesian or Jersey cows, while in the Austrian study, Austrian Simmental, Holstein Friesian and Brown Swiss are all plausible breeds.

	Kofler *et al.*, 2013	Whay *et al.*, 1998	Murray *et al.*, 1996	Kossaibati and Esslemont, 2000	Warnick *et al.*, 2001	Cook *et al.*, 2004
Country or region	Austria	UK	UK	UK	USA (NY)	USA (WI)
Breed	Unmentioned	Hol	Unmentioned	Hol	Unmentioned	Unmentioned
Number of dairy cows or herds	281 cows	53 lame (95 total) cows	37 herds	50 herds	2 herds	10 herds
Herd type	Varied	Not mentioned	Free-stall / grazing	Free-stall / grazing	Free stall / housed all year	6 free stall + 4 tie stall / housed all year
Type of measure	% of total animals with lesion	% of total animals with lesion	% of total lame diagnosis	% of total lame diagnosis	% of lame diagnosis (mean of 2 farms)	% of lame diagnosis (mean of 10 farms)
Laminitis	**67.5**	3.8	/	/	/	/
Heel horn erosion	**53.3**	9.4	4.0	1.0	/	1.0
White line disease	**37.8**	**26.4**	**22.0**	9.0	9.0	**10.0**
Sole haemmorhage	**29.9**	5.7	8.0	2.0	/	6.0
Horn fissure	**19.1**	/	/	/	/	/
Double sole	8.1	/	2.0	/	/	/
Limax	6.2	/	5.0	2.0	/	0.5
Corkscrew claws	4.7	/	/		/	/
Sole ulcer	3.8	**22.6**	**28.0**	**20.0**	**19.0**	**18.0**
Dermatitis digitalis	1.3	5.7	8.0	**20.0**	**32.0**	**57.0**
Phlegmon	/	**13.2**	5.0	**13.0**	**11.0**	1.0
Sole penetration	/	3.8	/	/	/	/

Appendix 2

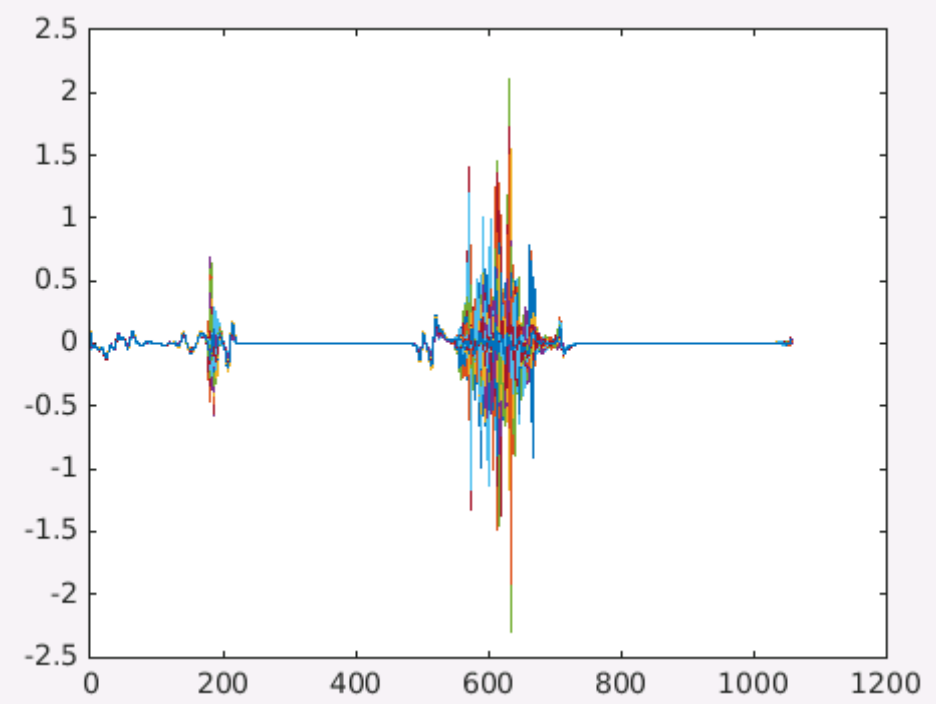

Figure 26: Original standardized MIR spectra.

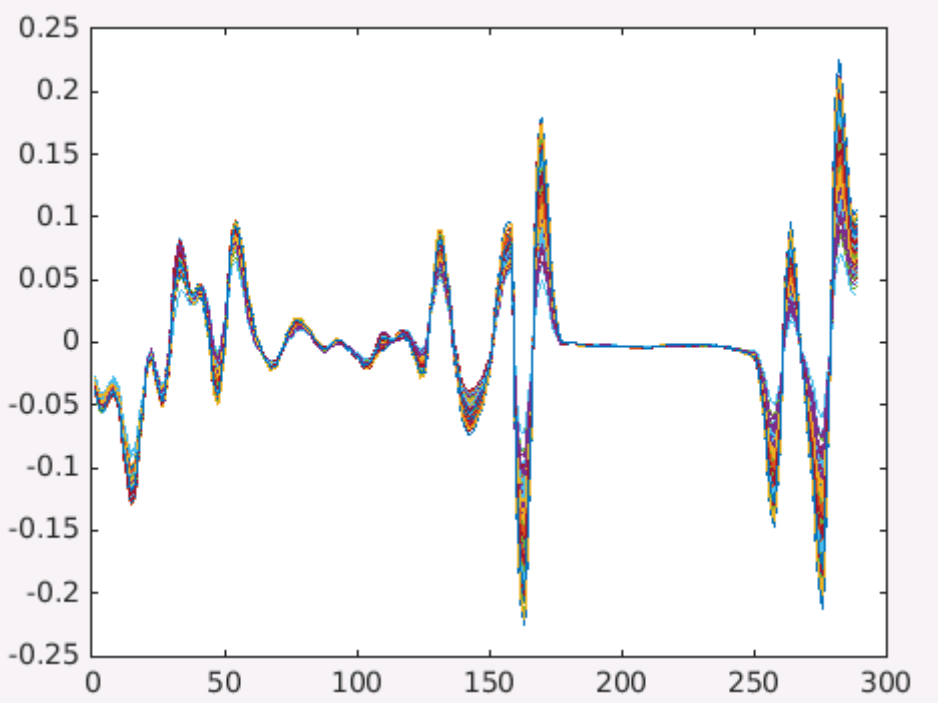

Figure 27: MIR spectra after first derivative as preprocessing.

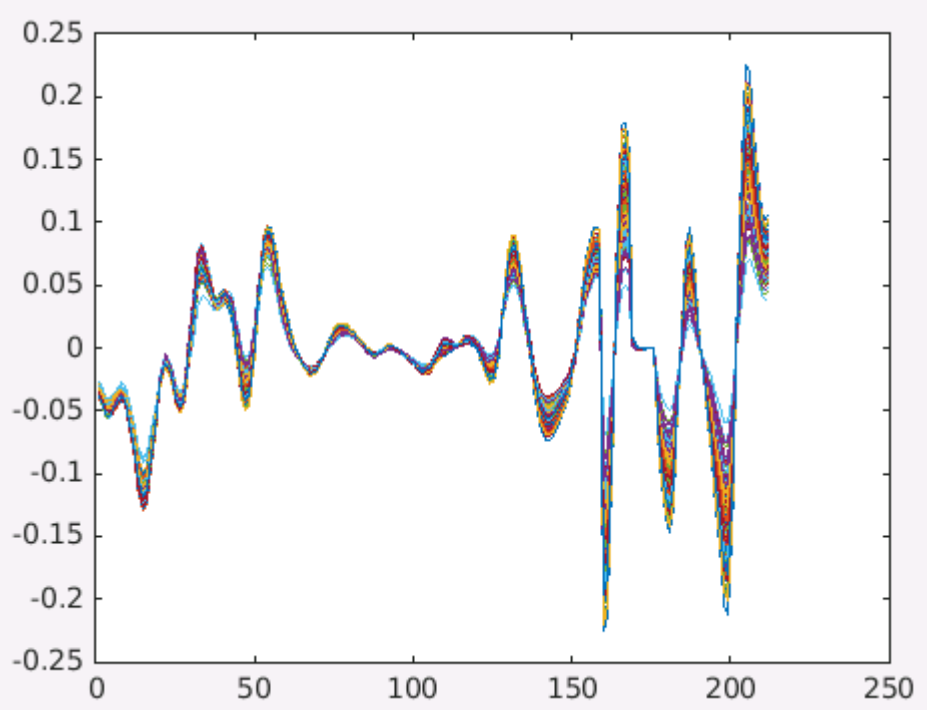

Figure 28: MIR spectra after selection of relevant wavenumber.

Appendix 3

Cow 1 (FL)	Months in milk									
5th parity	1	2	3	4	5	6	7	8	9	10
Locomotion score	3		3	2	5	4				
F:P ratio	1.41		1.25	1.03	1.78	1.38				
BHB	160.88		127.05	97.21	152.08	132.68				
Acetone	0.06		0.04	0.06	0.06	0.04				
Citrate	10.77		10.09	9.11	10.39	10.41				
C18:1cis9	1.17		0.81	0.63	1.37	0.85				
Ca	1323.15		1285.54	1270.62	1245.90	1241.27				
Na	384.17		388.28	393.53	343.44	442.42				
P	1152.41		1146.43	1193.45	1098.54	1100.73				
K	1552.28		1561.35	1571.52	1341.37	1465.50				
Hoof trimmer data						no lesion				

Cow 2 (BV)	Months in milk									
5th parity	1	2	3	4	5	6	7	8	9	10
Locomotion score			1	2	4	3	5			
F:P ratio			1.48	1.32	1.15	0.69	1.36			
BHB			144.05	162.17	131.15	118.80	206.83			
Acetone			0.07	0.09	0.05	0.03	0.03			
Citrate			7.73	9.45	6.71	6.26	6.56			
C18:1cis9			1.09	1.00	0.86	0.61	1.12			
Ca			1097.57	1137.73	1122.11	1132.10	1242.98			
Na			340.94	321.89	379.07	489.01	596.91			
P			916.65	933.84	1008.76	1015.48	1055.89			
K			1476.08	1572.27	1529.06	1529.42	1449.80			
Hoof trimmer data			no lesion			1 sole ulcer				

Appendix 4

The results obtained for the same data sets pre-treated with only a first derivative and both first derivative and SNV show that for this data, SNV is not necessary as the results are very similar.

PREDICTION						Calibration		Validation	
Subset	Pre-treatment	N	L	F	LV	Sensitivity (%)	Specificity (%)	Sensitivity (%)	Specificity (%)
All	1st derivative	7331	592	115	15	0.72	0.62	0.56	0.62
	1st derivative + SNV				15	0.7	0.62	0.58	0.63
First 2 months of lactation	1st derivative	1526	138	113	15	0.67	0.68	0.38	0.68
	1st derivative + SNV				15	0.66	0.69	0.35	0.71
Last 8 months of lactation	1st derivative	5805	454	115	15	0.74	0.62	0.63	0.58
	1st derivative + SNV				15	0.73	0.63	0.61	0.6

Appendix 5

		Locomotion score 1		Locomotion score 2		Locomotion score 3		Locomotion score 4		Locomotion score 5	
	Unit	Mean	St d	Mean	St d	Mean	St d	Mean	St d	Mean	St d
F:P ratio		1.28	0.21	1.291	0.23	1.29	0.23	1.33	0.26	1.35	0.33
BHB	µmol/l	167.26	49.20	160.07	59.22	157.87	47.59	163.79	54.79	163.59	55.12
Acetone	µmol/l	0.06	0.03	0.06	0.04	0.06	0.03	0.06	0.05	0.06	0.04
Citrate	mmol/l	9.45	1.58	9.14	1.59	9.05	1.66	9.07	1.76	9.18	2.21
C18:1cis9	g/100ml	0.82	0.21	0.82	0.22	0.84	0.23	0.89	0.26	0.93	0.36
LCFA	g/100ml	1.68	0.39	1.69	0.40	1.70	0.41	1.78	0.46	1.86	0.63
MCFA	g/100ml	2.24	0.45	2.25	0.48	2.19	0.42	2.22	0.45	2.23	0.56
SCFA	g/100ml	0.37	0.07	0.38	0.08	0.37	0.07	0.37	0.07	0.38	0.08
Ca	mg/kg	1202.42	98.01	1191.74	101.42	1184.85	96.58	1182.86	109.61	1187.48	107.98
Na	mg/kg	352.78	57.28	1019.07	94.33	1015.18	92.38	1006.05	90.17	1021.96	86.12
P	mg/kg	1028.41	90.44	1517.17	96.97	1521.97	88.82	1506.33	98.36	1497.91	109.55
K	mg/kg	1534.64	92.45	358.26	56.71	363.65	58.35	363.64	61.84	361.54	55.61