KU LEUVEN

FACULTEIT BIO-INGENIEURSWETENSCHAPPEN

Investigation of the potential use of walking speed measurements in automated lameness detection systems for dairy cattle

Onderzoek naar het potentieel van loopsnelheid in automatische systemen voor kreupelheidsdetectie bij melkvee

Promotoren:
Prof. Daniel Berckmans
Dr. Claudia Bahr
Departement Biosystemen
Afdeling M3-BIORES: Meet, Modelleer &Manage Bioresponsies

Masterproef voorgedragen
tot het behalen van het diploma van
Master of science in de bio-ingenieurswetenschappen:
biosysteemtechniek

Jens Bonnast

juni 2015

Preface

This thesis project was supported by the faculty of Bioscience Engineering of the KU Leuven. The author of this study wants to thank the M3-BIORES lab for the thesis subject and for the use of their facilities and infrastructure. The author also wants to thank the owner of the farm in Arendonk, Peter Blockx, and his son for the use of the farm facilities and the cows. Their help during the experiment was appreciated very much.

Two people have been indispensable to achieve the results presented in this study. Without their expertise this work would have looked completely different. The author wants to thank Claudia Bahr for her advice during the project and for her feedback. She also stressed the importance to keep it clear and comprehensible which was appreciated very much. The help and advice of Tom Van Hertem, in every aspect of the thesis project, were also irreplaceable. He was always willing to help if problems occurred and his software related tips were very useful. The author also wants to thank professor Daniel Berckmans for his attendance and feedback during the mid-term presentation.

Abstract

Lameness detection is an important problem on modern dairy farms. Lameness negatively influences animal welfare and health. It also causes economic losses for the farmer. Due to focus on intensification and scaling up, the prevalence of lameness has increased. To reduce the negative effects, an early detection is important. Manual lameness detection systems exist but they are time-consuming and their performance is limited. Automated lameness detection models (ALDM) have been developed to solve this problem and to give an automatic alert of the onset of lameness as soon as possible. The number of variables used in ALDM is large, but the use of walking speed has not been investigated yet.

In this study the impact of lameness on the walking speed of dairy cattle was investigated and the effect of walking speed on the performance of ALDM was determined. In 17 ALDM that were analysed, back posture and activity related variables were frequently used. In this study those variables were combined with walking speed in the development of three types of ALDM: statistical models, dynamic autoregressive models (dAR) and dAR models with exogenous variables.

Data were collected on a commercial dairy farm in Arendonk, Belgium. An experiment was set up to collect manual and automated walking speed measurements. These walking speed measurements were normalised and analysed using an ANOVA repeated measures analysis to investigate the effect of lameness on the walking speed of the cows. For the development of the ALDM only automated measurements were used. Model structures were optimised using the Akaike information criterion. Different combinations of the variables were made and the performance of the models was determined using ROC curves. The AUC values were calculated for each model and the most promising models were selected and validated.

The highest specificity was equal to 81,6% and was found for a statistical model using back posture and walking speed information. The sensitivity of this model was equal to 18,3%. In general the specificity was higher for models using the walking speed information but the difference compared to models without walking speed information was smaller than 5%. The performance of the dAR and dARX models was not higher than the performance of the statistical models due to practical limitations restricting the potential use of dynamic information of the measured variables.

Keywords: animal welfare, lameness, dairy cattle, detection, walking speed

Abbreviations, tables and figures

List of abbreviations

ACT:	variable containing the average daily activity over the last week [steps/day]
AIC:	Akaike information criterion
AMS:	automatic milking system
ANOVA:	analysis of variance
AUC:	area under curve
ALDM:	automated lameness detection model(s)
AWS:	normalised automated walking speed
B_{thresh}:	threshold on b_0 parameters
BPM:	variable containing the average back posture over the last week [dimensionless]
BPM / WS:	ratio of BPM over WS [s/pixel]
dAR:	dynamic autoregressive
dARX:	dynamic autoregressive with exogenous variables
DD:	digital dermatitis
GRF:	ground reaction forces
LMV:	limb movement variables
LS:	locomotion score
MLDS:	manual lameness detection system(s)
MW:	moving window
MWS:	normalised manual walking speed
P:	lameness probability
P_{thresh}:	lameness probability threshold
ROC:	receiver operating characteristic
SU:	sole haemorrhage and ulceration
WLD:	white line disease
WS:	variable containing average walking speed over the last week [pixels/s]

List of tables

List of figures

Table of contents

1. Introduction

In this thesis project the potential use of walking speed as a new variable in automated lameness detection models (ALDM) for dairy cattle was investigated. The project was started with an extended literature study describing the current situation in the dairy sector. A background study was made and the problem of lameness was situated. An overview of lameness detection models available nowadays was given. The different models were discussed and the gap in literature was defined.

An experimental protocol was developed to investigate the effect of lameness on the walking speed of dairy cows. Manual and automated walking speed measurements were collected on a commercial dairy farm and analysed.

For the development of ALDM, measurements for three variables (activity, back posture and walking speed) were collected over time. The measurements were processed by three different modelling strategies: statistical modelling, autoregressive modelling and autoregressive modelling with exogenous variables. Model structures were optimised. The performance of the developed models based on activity and back posture was determined and the best models were selected. New models based on combinations of all three variables were developed and their performance was also determined. The best models from this group were selected as well. The specificity, the ability to detect non-lame cows, and the sensitivity, the ability to detect lame cows, of the selected models were compared to investigate the effect of the use of walking speed measurements on the performance of ALDM. Eventually, the best classification model was identified.

2. State of the art

2.1 The dairy sector of today

In order to place the work presented in this study in its context, a sketch of the dairy sector in Flanders (regional), Belgium (national) and the European Union (international) was made.

Regarding Flanders, the number of dairy cattle has decreased with more than 14% from 327.067 cows in 2000 to 279.171 in 2012. In 1984, Flemish dairy farms housed 533.875 animals, almost twice as much cows as they do nowadays. The number of specialised dairy farms has dropped as well: from 4.303 farms in 2005 to 2.794 farms in 2012. The average number of cows per farm increased from 71 cows in 2005 to 100 cows in 2012. In the same period, total milk production increased with 20% from 1.784 million kilogram in 2004 to 2.144 million kilogram in 2013. This was achieved by increasing the individual milk production from approximately 6.500 to 7.800 kilogram per cow. These figures prove the two common trends in all agricultural sectors: farms grow bigger (scaling up) and production increases due to continuous improvements (intensification). For Flanders, the total production value of milk and milk derivatives, showed large variations in the last decade with a minimum of 459 million euro in 2009 and a maximal value of 719 million euro in 2011 (Platteau et al., 2012). This large variation is due to the large volatility of milk prices in the same period. The average milk price for the farmer lies around 0,30 euro per kilogram. Within a period of 20 months however, prices have varied between 0,41 euro per kilogram in October 2007 and 0,20 euro per kilogram in June 2006 (van Winsen et al., 2011).

The same trends are valid on a Belgian level: the number of dairy cows decreased from close to 600.000 cows in 2000 to 488.000 cows in 2011. Total milk production in Belgium increased slightly, from 3.055 million kilogram in 1985 to 3.215 million kilogram in 2010 (Versonnen, 2012).

If extended to the level of the European Union, similar evolutions have taken place. As the milk production per cow increased up to approximately 6.500 kilogram on average, the total dairy herd of all European Union members has been decreasing steadily to 23.475 million cows in 2013. Milk production in the European Union represented a total value of 53.100 million euro in 2011 which was equal to about 14% of the total agricultural output (Marquer, 2014).

All three the levels discussed above show the same trends: the total number of cows and the number of specialised farms decreases whereas the production per cow and the average size of individual farms increases. These trends are known as scaling up and intensification and are widespread in all agricultural sectors. Prices in the dairy sector are relatively low and volatile.

2.2 Challenges in the dairy production

In the past, farms were smaller and farmers had a close contact with their animals allowing them to detect problems in early stages. In the last decades, due to focus on intensification and economic productivity, the number of health and welfare problems has increased. Lameness for example is a major concern in the dairy industry causing welfare problems and economic losses. Since time per animal is limited on modern dairy farms, all kinds of systems have been developed and integrated on farms to monitor the animals and to facilitate the work of the farmer. On commercial dairy farms, manual milking is replaced by mechanised milking robots (Borderas et al., 2008) and automated systems using sensor data are used to predict the moment of calving or the moment of heat to optimise the time of insemination (Williamson et al., 2006). Other aspects however are not completely automated yet.

2.3 Definition of lameness and its appearance

Foot and claw problems in cows are a widespread problem in dairy cattle. In early stages, the effects of these problems are limited (Flower and Weary, 2009) but if not treated properly the locomotion of the cows aggravates. This new state is called lameness and is often referred to as a deviation in gait due to physical pain as a result from any type of claw or leg injuries or disease (Flower and Weary, 2009). Archer et al. (2010a) use a similar definition where the term lameness is used to describe a clinical representation of impaired locomotion and mobility regardless the cause.

Both claw problems and the resulting lameness state are common problems on dairy farms all over the world. Authors in different research facilities over the world have been working on models to facilitate lameness detection. Claw problems affect up to 96,7% of all cows (Cramer et al., 2008). Although a monthly study performed in the United Kingdom between August 2008 and July 2009 proved that more than 93% of the cows was found to be lame at least once during the covered period (Archer et al., 2010b), the mean annual prevalence of lameness lies lower but still reaches a value of approximately 20% (Clarkson et al., 1996).

2.4 Impact of lameness on dairy cattle

Foot disorders and the resulting lameness state affect both the animals and the farmers. Especially in the beginning, consequences are very limited and cow behaviour is hardly influenced. This is, at least partially, due to the stoic nature of cattle: even painful injuries at claws or legs may cause little or no clear behavioural effect (Flower and Weary, 2009). In an evolutionary context cows are prey animals and pain and weakness are kept hidden as long as possible (Weary et al., 2006). Later on, in more painful stages, the effects become more visible because gait (Pastell and Kujala, 2007), posture (Poursaberi et al., 2010) and general behaviour (de Mol et al., 2013) of affected cows alter. Lame cows often suffer pain of long duration what negatively influences the general welfare (Alban, 1995, Whay et al., 1998).

2.5 Economic consequences for the farmer

The effects of lameness are also considerable for the farmers. Lameness causes reproductive problems and is generally considered to be one of the main health problems causing economic losses for dairy farmers (Enting et al., 1997). Nowadays, farmer awareness of the problem is very low. According to Fabian et al. (2014) only 27,3% of the cows with mobility problems was identified correctly by the farmers. By increasing the awareness of the economic consequences related to lameness, farmers might be more eager to take action what could have a positive influence on animal welfare as well (Bruijnis et al., 2010).

2.5.1 Milk production

The first economic factor affected by lameness is milk production. From earlier studies the impact of lameness state on milk production was not clear. The results varied from a decrease in milk production (Rajala-Schultz et al., 1999) over no effect at all (Martin et al., 1982) to a higher milk production for lame cows (Dohoo and Martin, 1984).

More recent studies, that can be assumed to be representative for modern farm management techniques, show a more consistent decreasing trend. Lame cows show a significant decrease in milk production compared to their healthy herd mates. Milk production was around 2,6 kg/day lower for lame cows (Warnick et al., 2001). This result is similar to the estimated production losses of 1,5 to 2,8 kg/day two weeks after lameness diagnosis (Rajala-Schultz et al., 1999). If the relationship between milk production and lameness is regarded over an entire 305 days lasting lactation period, lameness leads to a decrease in milk production of 424 kg per cow (Bicalho et al., 2008).

High-yielding dairy cows are more susceptible to diseases and other problems due to a negative energy balance. The energy balance is defined as the difference between energy intake (feed) and the total amount of energy needed for maintenance and milk production (Fenwick et al., 2008). According to Bauman and Currie (1980) the energy balance of high-yielding cows turns negative around calving. The higher the milk production, the more negative the energy balance and the longer it takes before the energy balance returns to a positive state. Milk yields per cow have increased drastically and are expected to increase further in the near future so lameness is likely to become an even greater problem if no action is taken (Archer et al., 2010b).

2.5.2 Cow reproduction

Lameness has an effect on cow reproduction as well (Bruijnis et al., 2010). The impact of lameness on cow reproduction can be investigated by looking at the ovarian cycle and a possible delay in this cycle. In a study conducted on a Belgian dairy farm, 20,5% of the cows showed a delayed resumption of the ovarian cycle in the first 60 days postpartum (Opsomer et al., 1998). Another study, involving 536 Holstein cows and performed between June 2002 and May 2003, showed the relationship between degree of lameness and delayed cyclicality. The cycle was said to be delayed if progesterone levels were consistently lower than 1 ng/ml blood in the first 60 days postpartum. According to this definition, 6% of the non-lame cows showed a delay in their ovarian activities compared to 17% of the severely lame cows. Moderately lame cows had an intermediate result with delayed ovarian cycles in 14% of the cases (Garbarino et al., 2004). It has also been proven that lame cows are less likely to show any type of oestrous behaviour compared to their non-lame herdmates. Non-lame cows express oestrous behaviour 2,8% (±0.6) of the time compared to 1,8% (±0.4) of the time for lame cows.

2.5.3 Culling

A third direct effect of lameness resulting in an economic loss is culling. Lameness is the third reason for culling, following infertility and mastitis (Junge, 1997). The effect of lameness on culling depends whether lameness is detected in the beginning or at the end of a lactation. Cows that are diagnosed lame in the first half of the current lactation period, are two times more likely to be culled (Booth et al., 2004). If lameness occurs in a later phase of the lactation, the association with culling is less clear. This can be due to the fact that the cow is confirmed pregnant again already or the impact on total milk yield is less important (Booth et al., 2004).

The factors above represent direct production losses due to lameness. Lameness also causes a series of costs related to the treatment of the problem such as visits of the veterinarian or the claw trimmer. If all these costs are brought into account, lameness may lead to an annual loss of 75 US dollar or approximately 50 euro per cow per year (Bruijnis et al., 2010). If extrapolated to an entire farm of 200 animals, economic losses related to lameness are considerable (10.000 euro per year). It is in the interest of both the animals and the farmers that lameness is detected in an early stage. By doing so, treatment is relatively easy and economic losses stay limited (Bruijnis et al., 2010).

2.6 Causes and risk factors of lameness

2.6.1 Causes

Rather than a disease, lameness is a clinical sign associated with a broad range of conditions (Potterton et al., 2012). Although other causes exist, the majority of lameness cases is related to the presence of four different foot diseases (Cramer et al., 2008). Those diseases are sole haemorrhage and ulceration (SU), white line disease (WLD), digital dermatitis (DD) and interdigital necrobacillosis. For a description of those, as well as the less common diseases, the work of Cramer et al. (2008) can be consulted.

Cramer et al. (2008) also state that a distinction between cubicle housing systems and tie-stalls has to be made if the presence of claw diseases is discussed. Cubicle housing systems, are known to increase the incidence of lameness relatively to tie-stalls. The lowest percentage of lameness in cubicle housing systems (19%) was found in cubicle stalls with a straw bedding (Rouha-Mulleder et al., 2009). Despite the importance of the housing type, a few general trends are reflected. In both the housing types, infectious lesions are the most common. Cows in free-stall systems (46,4%) are affected more often than cows in tie-stalls (25,7%). DD is the most widespread problem affecting 9,3% of the cows and 69,7% of the herds in tie-stalls. The prevalence of DD for free-stalls is even higher affecting 22,7% of the cows and 96,7% of the herds. The most common types of claw horn lesions are haemorrhages and ulcers. 7,1% of the cows and 4,7% of the herds in tie-stalls show this type of lesion. The values in free-stalls are higher reaching 11,0% and 9,2% respectively (Cramer et al., 2008). As DD is the most common problem, its treatment is described properly. Knowledge of SU and WLD treatments is more limited (Potterton et al., 2012).

2.6.2 Risk factors

Factors associated with an increased lameness prevalence can be divided into different groups. First of all the breed is an important aspect. Herds consisting of Holstein-Friesian cows have a lameness prevalence between 38% and 45%. Herds consisting of one or more other breeds show lower lameness prevalence values around 15% (Barker et al., 2010). Other factors can be subdivided in environmental and management related factors (Cramer et al., 2008).

2.6.2.1 Environmental factors

A number of environmental factors within the housing and grazing areas is known to increase the prevalence of lameness on a farm (Barker et al., 2010). The housing system and the concomitant access to pastures are important. Housing systems for dairy cows differ in the period over which cows have access to outdoor pastures. In regions with sufficient space and an appropriate climate, cows are allowed to graze outdoors the entire year. In other climates cows are grazed outdoors in summer months only. Zero-grazing or continuous housing systems, in which the cows are housed throughout the year, are common as well. Especially in densely populated regions, where land is limited, grazing is not the most efficient or cost-effective use of the land (Haskell et al., 2006). In continuous housing systems, cows can be fed concentrate feed more easily. These systems are therefore often preferred on specialised farms housing high-yielding cows. The length of the housing period was found to affect the prevalence of lameness. Haskell et al (2006) proved that the prevalence of lameness in grazing herds stays limited to 15%. In zero-grazing herds however, lameness affects up to 39% of the cows.

Next to pasture access, the type of floor is an important environmental factor as well. According to Somers et al. (2005) cows on concrete floorings show claw disorders in about 80% of the cases. In straw-yard farms this value lies significantly lower at 55 to 60%. The negative effect of concrete is possibly linked with its unyielding nature (Haskell et al., 2006).

Other structural factors affecting the prevalence of lameness include the presence of lying spaces (Cook and Nordlund, 2009), the amount and the quality of bedding in free-stalls (Barker et al., 2007, Fregonesi et al., 2007) and the exposure to sludge which negatively influences the hardness of the claw horn. Softer claws are more sensitive to lesions and digital dermatitis (Borderas et al., 2004, Somers et al., 2005, Gregory et al., 2006).

2.6.2.2 Management related factors

Beside the environment, the management of the cattle within this environment is also important. Factors associated with cattle management are often more difficult to quantify. Although most aspects are subtle, a few important management related factors are known to influence the prevalence of lameness.

The first aspect is the management of the claw health. Claw overgrowth increases the risk of lameness in dairy cattle. Another aspect related to claw growth is the use of claw trimming as a preventive treatment for lameness (Klaas et al., 2003). Aoki et al. (2006) proved the positive influence of claw trimming on walking characteristics such as step length.

The use of automatic scrapers increases the risk of lameness by 9%. In farms were no automatic scrapers were used, Barker et al. (2010) found a lameness prevalence of 35,5%. In farms were those automatic scrapers were used, lameness prevalence was equal to 44,5%. The same study also gives a possible explanation for this relationship between the use of automatic scrapers and lameness prevalence: the movement of the scraper disturbs the cows and forces them to move out of its way. In places were cows are close to each other they are not always able to see the scraper coming and are therefore forced to make hurried movements. These sudden movements increase the risk of claw injuries since cows do not have the time to place their claws properly. Physical damage caused by direct contact between the cow and the scraper is also possible (Barker et al., 2010).

Which of the causes and risk factors is dominant varies between different countries, regions and even between different farms (Manske et al., 2002, Tadich et al., 2010, Leach et al., 2012). Differences in breeds, feeding and general management strategies between different farms in different regions or countries are a possible declaration for this local character.

A vital part of lameness control on farms consists of a rapid and effective treatment (Leach et al., 2012), reducing the number of lame cows and reducing the prevalence of lameness in a sustainable way. Therefore the development of models allowing an early identification and effective treatment is crucial. In combination with the implementation of preventive measures these models could improve the current situation (Potterton et al., 2012).

2.7 Lameness detection

Treating lameness, especially if detection in the early stages is aspired, requires valid and reliable methods for the detection of cows with claw problems and lesions (Flower et al., 2006). Since the last decade of the 20th century, several different systems have been developed to analyse the locomotion of cows. Those systems are called locomotion scoring systems and are split up into 2 main groups: manual locomotion scoring systems and automated locomotion scoring systems. The recent importance of locomotion scoring systems is illustrated by the fact that 70% of all articles regarding those systems was published after 2007 (Schlageter-Tello et al., 2014). This present interest can be declared by the association between increased milk yield and the incidence and prevalence of lameness. Depending on the specific system used in a particular case, each locomotion score corresponds with a certain degree of impaired locomotion. If the score exceeds a critical value, that is once more depending on the used scoring system, a cow can be identified as lame (Schlageter-Tello et al., 2014). Since locomotion scoring systems can be used to detect lame cows, they are also called lameness detection systems. This name will be used in the scope of this text.

Flower and Weary (2009) assume that the claw injuries and diseases discussed above can affect the gait of cattle. Although that not every injury will directly influence the gait, this relationship is used as a basis to detect problems. In order to increase the ability to identify problems, the effects of shape and posture should be controlled for. Tall cows with long legs for example are expected to have longer strides than smaller cows (Flower and Weary, 2009).

2.7.1 Manual lameness detection systems

The first systems were all manual lameness detection systems (MLDS) and were developed to give a good indication of the locomotion quality of the cows (Schlageter-Tello et al., 2014). In those systems human observers evaluate certain gait and posture characteristics to score the locomotion. The scores on the different characteristics are combined to achieve a general view of the level of impaired locomotion. Gait characteristics are associated with alterations of the limbs. The asymmetrical character of the gait and the reluctance to bear weight are two of the most commonly used gait variables. Posture characteristics are related to alterations in body parts other than the limbs. Examples of this group of characteristics are head bobbing and curvature of the back (Schlageter-Tello et al., 2014).

In the last decades, research regarding lameness in dairy cattle has increased considerably (Clarkson et al., 1996). Since 1988, 25 different MLDS have been discussed in literature. Although that preferences may differ regionally and among different researchers, some of those systems are more popular than others. The MLDS described by Sprecher et al. (1997) was mentioned in about 28% of the articles discussing lameness. The methods described by Manson and Leaver (1988) or Flower and Weary (2006) were also popular. A combination of those three systems was often used as a basis for other MLDS. An overview of all 25 the MLDS developed since 1988 is given in the review presented by Schlageter et al. (2014).

2.7.1.1 Advantages MLDS

The use of MLDS has some important advantages. MLDS do not require any type of specialised equipment what allows an easy applicability under farm conditions. Those manual systems are inexpensive and non-invasive as well, two important characteristics in agricultural applications. The observer also gets the opportunity to look at all the individuals in the herd at once and the results are known immediately (Renn et al., 2014). The scoring can even be performed by the farmer himself (Whay, 2002).

2.7.1.2 Disadvantages MLDS

The fact that many different systems were developed and used, proves the lack of consensus. Each of the MLDS discussed in literature uses different scales and investigates lameness using a different combination of gait and posture characteristics. The number of possible characteristics is high. Regarding the scaling systems, a distinction between discrete and continuous scales can be made. This broad range of possibilities is a major disadvantage of manual systems. Results obtained using different systems are completely independent and can never be compared with each other. Other limitations are related to the dependence of an observer who requires a certain level of training. The use of human observations inevitably leads to subjective systems as discussed in the next sections.

The usefulness of MLDS is limited by their validity, the reliability and the sensitivity (Flower and Weary, 2006). The validity can be investigated by comparing the locomotion scores given to lame and non-lame cows respectively. In many cases the relation between locomotion scores and the presence of claw injuries or lesions is rather poor. The variation in gait scores was only declared for 20% to 70% by the presence of known claw injuries or lesions in previous studies (Whay et al., 1997, van Eerdenburg et al., 2003, Flower and Weary, 2006).

Manual locomotion scoring is a subjective technique and, as is often the case for subjective methods, the reliability can vary over time within an observer or between observers (Flower and Weary, 2006). The reliability of a system can be investigated in two different ways. First of all the agreement of the scores of the same observer on multiple occasions can be used as a reliability measure (Flower and Weary, 2009). Another method is to compare the scores of different observers. O’Callaghan et al. (2003) found out that a trained observer was consistent over two scoring sessions in only 56% of the observations. In the same study the scores of two different observers agreed in only 37% of the cases. Observers may affect the measurements in three different ways according to Hollenbeck (1978):

1) Influence of sources of bias

Errors may occur by omission or they might result from observer expectations. The first is for example the case if an observer fails to score a behaviour that occurred while the latter can happen if an observer does not score a cow as lame if the rest of the herd is in good condition (Flower and Weary, 2009).

2) Experience of the observer

The effect of experience was illustrated by Main et al. (2000) in a study concerning lameness in pigs. The agreement between scores from experienced and inexperienced observers varied between 26% and 53%. Agreement between scores from experienced observers was much higher reaching values of 94%. A study performed by March et al. (2007) reported that at least 200 to 300 cows are needed to train an observer.

3) Observer drift

Observer scores can change over time (Hollenbeck, 1978). This is called observer drift and is especially problematic for live observations since scoring can be performed only once leaving no room for quantifying drift by re-scoring (Flower and Weary, 2009).

Next to the observer, the MLDS used in a specific situation may affect the results as well. Some MLDS use very general terms. Those systems will automatically lead to smaller differences between different observers. The more specific the terminology and the more detailed the system, the bigger the differences between different observers (Flower and Weary, 2009). A continuous scale allows observers to notice more subtle changes in behaviour (Flower and Weary, 2006).

The last limiting factor is the sensitivity of the MLDS. In this context the sensitivity has to be seen as a measure for the ability of the system to detect lame cows. The use of systems with a continuous scale proved to be more sensitive than discrete scoring scales for lameness studies in sheep (Welsh et al., 1993).

As mentioned earlier, cattle is known to have a stoic nature. In combination with the aspects discussed above, this stoic nature may declare why it is often difficult to judge changes relative to ‘normal gait’ correctly. The number of cows with claw injuries or disease that is correctly identified using MLDS varies between 25% (Whay et al., 2003) and 33% (Espejo et al., 2006). These values are often insufficient to allow a reliable detection of claw problems in an early stage.

2.7.1.3 Conclusion MLDS

MLDS are easy to use under practical conditions but they are time-consuming and subjective techniques for assessing lameness in dairy cows. They show a lack of clear standards and depend on the skills of the observer to detect subtle changes in locomotion (Winckler and Willen, 2001). In combination with high time pressure on modern farms and increasing farm sizes, which inevitably leads to less time per animal, these factors are the main reason why manual systems are more and more replaced by automated systems where possible (Diskin and Sreenan, 2000). Other important aspects favouring automated systems are high costs of manual labour and thin profit margins forcing farmers to seek efficiencies in all aspects of their business (St-Pierre, 2001).

2.7.2 Automated lameness detection models

Automated lameness detection models (ALDM) have been developed to give the farmer an automatic alert of the onset of lameness as soon as possible (Alsaaod et al., 2012). In the review presented by Schlageter-Tello et al. (2014), 15 different ALDM are listed. Although the review was published very recently, the most recent studies, such as the ALDM based on consecutive 3D-video recordings by Van Hertem et al. (2014), are not included. This proves that automated lameness detection is still an ongoing hot topic. In the overview a distinction between three different approaches is made: the kinetic approach, the kinematic approach and the indirect approach. The different approaches are used in 33%, 37% and 30% of the cases respectively (Schlageter-Tello et al., 2014) and are discussed in the next sections. Table 1 and Table 2 summarise the applied approach, sensor type(s) and measured variable(s) that are used in the ALDM published in literature since 2002.

2.7.2.1 ALDM using the kinetic approach

Hall (1995) defined kinetics as the study of forces involved in motion. ALDM following the kinetic approach typically use force plates or weight recording units to collect data (Flower and Weary, 2009). Both these methods measure the forces when the claw contacts the floor. Scott (1988) and Van der Tol (2002) have used these methods to measure forces exerted by walking dairy cows.

The model developed by Rajkondawar et al. (2002b) is an ALDM based on measurements of ground reaction forces (GRF) exerted when the claw contacts the ground. Measurements are obtained from two parallel-force plates accommodated with side railings to guide the cows. These railings guarantee that the left limbs contact the left force plate and the right limbs the right force plate (Rajkondawar et al., 2006).

A kinetic approach requires a thorough knowledge of the walking pattern of cattle that can be divided in a weight-bearing, starting when the claw contacts the ground and lasting till the moment the foot is lifted off the ground, and a non-weight-bearing phase. During the weight-bearing phase GRF are transmitted by the foot to the ground. These forces can be measured over time and can be used for the calculation of limb movement variables (LMV). According to Scott (1989) lameness influences the measured forces. In the model developed by Rajkondawar et al. (2006) this reasoning is the basis for the use of LMV in an ALDM. A correlation between the measured forces and the locomotion scores of the cows involved in the study was found. How the force measures are actually related to claw or leg injuries is less clear but Rajkondawar et al. (2006) conclude that their method allows an automated and accurate detection of lameness in dairy cattle. This model was used in the commercially available StepMetrix™ system (BouMatic, Madison, WI). Bicalho et al. (2007) tested the system under practical conditions and calculated a lower sensitivity (33,3% versus 67,5%) compared to manual locomotion scoring, performed by trained veterinarians. The specificity value of the StepMetrix™ system (89,5%) was higher than the specificity obtained from manual locomotion scoring (84,6%).

A second model based on the kinetic approach is the 4-balance system developed by Pastell et al. (2006). In this model changes in weight distribution were measured using four independent weight recording units. Although the model was able to detect problems, for example as a result from changes in leg loads, the specificity was too low. This problem was solved by Pastell and Kujala (2007) by using a probabilistic neural network. The final model reached an overall classification accuracy of 96,2% with a lameness detection rate of 100%.

ALDM based on the kinetic approach have proven to achieve good results. Practical limitations of ALDM based on the kinetic approach are related to the positioning of the claws on the weight recording units (Pastell and Kujala, 2007). The high costs of the used components are a major disadvantage of these models. Pressure sensitive materials such as force plates are expensive and have to be replaced regularly due to damages caused by direct contact with the claws. As mentioned above profit margins in the dairy sector are small, and costs, including the costs related to lameness detection, have to be kept as low as possible.

2.7.2.2 ALDM using the kinematic approach

The second type of ALDM relies on the kinematics of the cow movement. Kinematics studies the position changes of body segments over time (Hall, 1995). Kinematic procedures have a longer history in sports and equine applications, but they have been adapted for applications regarding dairy cattle as well.

In lameness detection applications, measurements are obtained from small spheres or markers attached to specific anatomical locations of the body of the cow. Popular locations include claws, limb joints and the back-line contour (Schlageter-Tello et al., 2014). The movement of the spheres is captured and the recordings are analysed using motion analysis software. This software calculates both linear and angular displacements as well as velocities and accelerations of the markers (Peham et al., 2001). Schlageter-Tello et al. (2014) divided the ALDM using a kinematic approach into four different categories.

1) Direct video recordings

A first technique directly analyses video recordings of cows equipped with the needed markers to obtain the kinematic variables. Flower et al. (2005) were the first to use computer-aided kinematic techniques to evaluate dairy cattle movement. They extracted six variables (stride length, stride height, stride duration, stance duration, swing duration and claw speed) from the recordings. They proved that kinematic measures could be used to distinct cows having sole ulcers from healthy cows: a significant difference for stride length, stride height, stride duration and claw speed was found. The claw speed of healthy cows was higher than the claw speed of cows having sole ulcers: 1,11 ± 0,03 m/s versus 0,90 ± 0,05 m/s. Another clear difference was found in the percentage of triple support (the time the cow is supported by three legs) of cows having sole ulcers (42%) and healthy cows (18%). Few differences between cows having sole lesions and healthy cows were found.

Similar methods were used by Aoki et al. (2006) to quantify the effects of claw trimming. Video recordings were collected and image analysis software was used to extract variables regarding the walking pattern of the cows. Blackie et al. (2011b) explored gait attributes often used in MLDS and evaluated them objectively using new technologies. Next to gait characteristics, joint flexion and spine posture were evaluated for cows having different lameness states. Effects on joint flexion and spine posture were limited in this study. Stride length on the other hand was clearly negatively affected by lameness.

2) Image pre-processing

In this technique the original video recordings are transformed into successive binary images. By doing so the detection of certain body parts facilitates (Schlageter-Tello et al., 2014) and so does the processing. Song et al. (2008) were the first to explore the possibilities of image pre-processing in lameness applications in dairy cattle. They managed to capture the claw positions with good precision and found a linear relationship between trackway overlap, defined as the position of the hind claw compared to the position of the front claw on the same body side (O'Callaghan et al., 2003), and manually assigned lameness scores.

Poursaberi et al. (2010) were the first facing the problem of similarity between background and body colour of the cows. Additional measures were required to distinguish the cow from the background. Van Hertem et al. (2013a) illustrated the importance of a good distinction between animals and background if side profiles are preferred as is the case for monitoring gait profiles. Cows can be extracted from the background, that is often dynamic, using segmentation algorithms (Van Hertem et al., 2013a). The importance of an individual approach was illustrated by Viazzi et al. (2013). Since differences exist in the walking pattern of individual cows (Pluk et al., 2012), the effects of lameness are not identical for all the cows. To bring this large variation in variables into account, an individual body movement pattern score using back posture of every individual cow was used to classify the cows into 3 lameness classes. Sensitivity and specificity values were equal to 76% and 91% respectively.

Image processing in side-view 2D space has proven to be problematic in practical on-farm conditions. The positioning of the camera is not always possible and the dynamics of the background are a problem as well (Van Hertem et al., 2013a). This can be solved using a 3D camera in top view perspective but the field of view of this type of setup is limited. Viazzi et al. (2014) extracted information regarding the back posture of the cows from video recordings obtained from a 3D camera to develop a lameness detection model. The final model had an accuracy of 90% (Viazzi et al., 2014) but only one measurement per cow was used. Van Hertem et al. (2014) optimised the model by using consecutive measurements to lower the number of false alarms (e.g. if a cow falls). The incorporation of consecutive measurements led to a decrease in classification rate: 81,2% of the cows was classified correctly. This value lies lower than the value in the original model (Viazzi et al., 2014) but since data were obtained on four different days, the daily setup of the camera can declare this difference (Van Hertem et al., 2014).

3) Pressure sensitive walkways

Another method to obtain kinematic variables uses pressure sensitive walkways consisting of several pressure sensors. Every time a cows passes through the walkway, new measurements are collected. Although forces can be measured using this type of sensors, force itself is not used as a variable and therefore it should be regarded as a kinematic model. This method differs from other kinematic models in that way that only variables resulting from claw-floor-interactions are measured. This solves the problems resulting from dynamic backgrounds and image processing, but it also lowers the number of variables that can be used. The curvature of the back for example cannot be measured directly (Maertens et al., 2011). Based on this approach Maertens et al. (2011) developed the GAITWISE system which reaches specificity values between 86% and 100%. Values for sensitivity vary between 76% and 90%.

In most studies investigating the impact of lameness on dairy cattle, variation in variables is considered on a daily or even weekly basis. In human applications, variations within one measurement are also considered important (Van Nuffel et al., 2013). Since some systems, like the GAITWISE system measure gait variables multiple times during one measurement, this within measurement variation could also be used in cattle applications. The research performed by Van Nuffel et al. (2013) confirmed the importance of within measurement variation as a predictive variable in early stages of lameness. A combination of a pressure sensitive walkway and video recordings, as described by Pluk et al. (2012), is also possible.

4) Accelerometers

Accelerometers attached to the limbs of the cows can be seen as a transition between the kinematic approach and the indirect approach discussed further on. Pastell et al. (2009) attached an accelerometer to each limb of the cows involved in the study and used these data to reconstruct the different stages in the walking pattern. Therefore the model is seen as a kinematic ALDM.

In kinematic studies derived variables (such as speed and stride overlap) have proven more useful than basic variables such as stride duration (Flower et al., 2007). The use of kinematic models, especially the direct video recordings and the image pre-processing techniques, have the important advantages of being non-invasive and avoiding unnecessary stressful contacts (Van Hertem et al., 2013a). In addition models using these techniques can be incorporated in existing farm systems relatively easy. The use of video recordings also implies that continuous information regarding all the animals can be obtained without placing a different sensor on each individual which makes these models relatively cheap (Viazzi et al., 2013). If markers are used to point out anatomical locations, skin displacements can limit the accuracy (Schlageter-Tello et al., 2014).

2.7.2.3 ALDM using the indirect approach

ALDM using the indirect approach are based on behavioural and production variables. The number of variables that can be measured automatically is large and still growing (Rutten et al., 2012). In this study the overview is limited to the most common variables. The equipment needed to measure them is also given and discussed. Subsequently the actual ALDM relying on certain combinations of those variables are listed.

Popular variables

The most popular behavioural variable is activity. Other commonly used variables such as milk yield, milk composition and feed intake are more production related. Equipment measuring these variables is available on most commercial dairy farms. Next to these general variables, body temperature, rumination activity and hormone concentrations can also be interesting.

1) Activity

Activity measurements can be used in different ways. Even within the dairy sector several applications, including the observation of locomotion behaviour and the prediction of the moment of calving, are known. The major application can be found in the detection of heat or oestrus. According to Williamson et al. (2006) cows in heat walk up to four times more than they do under normal conditions.

In most cases a baseline representing the normal activity level is established for each cow and changes in activity are measured relatively to this baseline. The two main technologies used to measure and monitor activity patterns are pedometers and accelerometers (Mackinson et al., 2013). Pedometers (AfiAct, Afimilk Ltd., Kibbutz Afikim, Israel) are devices measuring the number of steps taken by each cow. They are often attached to the leg of the cow. Activity data obtained from a pedometer are processed in the sensor and every time a cow passes a receiver, the data are sent to a computer in the farm where they are analysed using specialised software. Accelerometers (Select Detect^{TM}, Select Sires, Plain City, US; DeLaval^{TM}, DeLaval AB, Tumba, Sweden) are similar but more complicated as they measure movements in three independent directions: side-to-side, up and down and front to back. One of the major differences between the sensors is the sampling interval, varying between 1 minute (e.g. IceTag^{TM}, IceRobotics Ltd., Roslin, UK) and 2 hours (e.g. HR LD tag, SCR Engineers Ltd.,Netanya, Israel). Accelerometers are often attached to the collar of the cow. Recently, accelerometers in combination with other sensors measuring multiple variables at the same time have been developed. This is the case for the ALT-pedometer (Brehme et al., 2008) measuring not only activity but also lying time and body temperature. Another system (HR-Tag^{TM}, SCR Engineers Ltd., Netanya, Israel) measures activity and rumination (Mackinson et al., 2013). The ear-attached movement sensor developed by Bikker et al. (2014) is also interesting. They use a 3D accelerometer (SensOor, Agis Automatisering BV, Harmelen, the Netherlands) attached to the ear tags of each cow. The behaviour of the cows is then classified based on the movements of the ears (Bikker et al., 2014).

2) Lying time and lying bouts

Related to the activity of the cows is the time a cow spend lying down. The possibility for a cow to show normal lying behaviour under all conditions plays an important role in the production and the welfare of the cattle (Bewley et al., 2010). Lying behaviour is therefore an important factor regarding both the animal and the farmer. If a cow is deprived of adequate lying time, this has a negative impact on the welfare (Cooper et al., 2008). Insufficient possibilities or infrastructure may implicate physiological and behavioural signs of stress. Out of frustration the cow is expected to show abnormal behaviour if her needs regarding lying time cannot be fulfilled (Cooper et al., 2008). Another variable related to lying behaviour is the number of lying bouts and their duration (Ito et al., 2010).

According to Juarez et al. (2003) changes in lying behaviour are related to lameness. Originally lying behaviour was only observed visually or from video recordings (O'Driscoll et al., 2008) limiting its use due to subjectivity and time constraints. In the last couple of years however, activity monitoring sensors (TinyTag, Gemini Dataloggers Ltd., Chichester, UK; Pedometer+^{TM}, SAE Afikim, Israel) have been developed that allow direct automated measurements of lying behaviour of dairy cows.

3) Milk yield and composition

Dairy cows are milked on a daily basis. In conventional, non-robotic farms, common practice is to milk the cows twice per day although some farm managers opt for three or even more milking sessions. These moments can be used to obtain additional information regarding the health status of the cows and to detect emerging health problems (Huybrechts et al., 2014).

Different types of sensors can be used to achieve information. Milk flow meters (FloMaster, DeLaval AB, Tumba, Sweden; AfiFlo, Afimilk Ltd., Kibbutz Afikim, Israel) can be used for yield monitoring. Lame cows show a significant decrease in milk production compared to their healthy herd mates. Other types of sensors include milk conductivity sensors (e.g. AfiLab, Afimilk Ltd., Kibbutz Afikim, Israel) and can be used to achieve information regarding the composition (fat, protein, urea, etc.) of the milk. Next to yield and composition, the presence of specific substances such as hormones can be investigated as well.

4) Feed intake

Although roughage intake can be monitored, this kind of systems is often limited to research facilities. In continuous housing systems cows are often fed a certain amount of concentrate feed in addition to the roughage mixture. The feed intake of lame cows decreases. Concentrate feed has a higher nutritional value and influences milk yield and milk composition positively (Keady et al., 2001). Especially in high-yielding cows, the use of concentrate feed may be beneficial. In order to measure individual feed intake in production environments automated systems have been developed (e.g. GrowSafe Model 6000, GrowSafe Systems Ltd., Airdrie, Canada). This kind of systems allows continuous data acquisition regarding concentrate feed intake of individual cows.

5) Rumination

Rumination is known to be influenced by changes in the health status of the cattle. Two different types of sensors are used to monitor rumination. A first type measures ruminal activity by means of a bolus inside the animal. This type of sensors simultaneously measures body temperature and pH (e.g. Sentinel™, Kahne Limited, Auckland, New Zealand). These sensors are invasive and other sensors are often preferred. A first alternative system uses a sensor attached to the cow's collar (e.g. Hi-Tag, SCR Engineers Ltd., Netanya, Israel) and measures the acoustics of rumination. A collar with a noseband that measures the number of jaw movements (RumiWatchSystem, Itin + Hoch GmbH, Liestal, Switzerland) is also possible.

Existing ALDM using the indirect approach

ALDM following an indirect approach are based on a selection of the behavioural and production variables listed above. Behavioural information is obtained from accelerometers attached to the neck collar or a limb of each animal. Production variables are measured using other sensors that are already available on most modern farms. Since the needed technologies are already installed, the use of these variables in ALDM is relatively cheap. Different combinations of variables have led to the development of seven different ALDM using the indirect approach since 2008. Those are discussed in a chronological way below.

The first model was developed by Borderas et al. (2008) and was based on the principles of automatic milking systems (AMS). The implementation of AMS on dairy farms has a dual effect: milk yields are expected to increase thanks to higher milking frequencies and manual labour decreases. Lame cows are expected to visit the AMS less frequently (Borderas et al., 2008). For the analysis, the cows were divided into two groups according to the number of AMS visits. Only 4% of the cows in the high visiting group was found to be lame compared to 32% of the cows in the low visiting group.

Ito et al. (2010) developed a diagnostic tool entirely based on lying behaviour obtained from accelerometer data. The number, the duration of and the variation between the lying bouts were evaluated. According to Alsaaod et al. (2012) a sensitivity value of 72% was reached for this model. The specificity was equal to 81%. The next ALDM was developed by Chapinal et al. (2010) and used 3 different types of sensors. Weight recording platforms were used to measure deviations in weight applied to the rear legs, accelerometer data provided information regarding lying bouts and 2D video recordings were used to calculate walking speed. Lame cows were found to show longer lying bouts, lower walking speeds and a greater asymmetry in the weight applied to the rear legs.

Another model based on lying time, standing time and milk yield was constructed by Blackie et al. (2011a). The required information was obtained from accelerometers and milk meters. Once more lying times were significantly higher for lame cows.

De Mol et al. (2013) developed an ALDM using information obtained from accelerometers, a milking robot and a scale for feed weight. The accelerometers were used to determine if a cow was lying or standing. Milking robots were used to measure total milk yields and the scale provided information about concentrate left over. The final model developed by de Mol et al. (2013) achieved an overall sensitivity of 85,5%. The specificity was equal to 88,8%.

Three different types of sensors collecting measurements for five different variables were used in the ALDM developed by Kamphuis et al. (2013). Accelerometers and a weight scale provided information regarding activity and body weight respectively. Milk metres were used to measure the total milk yield, the total duration and the order in which cows were milked.

The sensitivity of this model varied between 40,1% and 56,5%. The specificity reached values between 80 and 90%.

The most recent ALDM using behavioural and production related information was developed by Van Hertem et al. (2013b). An accelerometer was attached to the collar of the cows providing neck activity and rumination measurements. Milk metres were used to determine the total milk yield of the cows involved in the experiment. A correct classification rate of 86% was found. Sensitivity and specificity were equal to 89% and 85% respectively (Van Hertem et al., 2013b).

2.7.2.4 Validation ALDM

Validation has to be approached critically. ALDM are often validated using the results of a manual locomotion scoring session as a gold standard. This implies that validation results are often influenced by the observer performing the scoring. Values of sensitivity and specificity as well as ROC-curves (the area under the curve) can be used as measures of validity. Most ALDM have high specificity values, exceeding 80%. According to Schlageter-Tello et al. (2014) the variation in sensitivity is large (39%-90%). In general the sensitivity of ALDM exceeds the sensitivity of MLDS. The specificity and the sensitivity of the ALDM suggest that non-lame cows are detected more accurately than lame cows.

2.7.2.5 Conclusion ALDM

Different types of ALDM have been developed. All of them have their specific characteristics, their advantages and limitations. If the main problems related to MLDS are compared with the situation where ALDM are used, it is clear that the use of ALDM solves the subjectivity factor. No observer is needed if automated models are used and all the factors related to the influence of an observer can be neglected. Time limitations and reliability problems are not as serious for ALDM either (Schlageter-Tello et al., 2014).

The possibilities for ALDM are big and every author has different preferences. An overview of all the automated models published in literature since 2002 as well as the approach and sensors they used is given in Table 1. The variables used by the different ALDM are divided in different categories (gait, posture, behaviour-production and speed) and are listed in Table 2. For an overview of the listed variables, including a definition of the used terminology, the work of Schlageter et al. (2014) can be consulted.

Table 1: Chronological overview of the different ALDM published in literature since 2002. The approach they are based on as well as the sensors they use are presented for each ALDM.

Model	**Approach**			**Sensor types**									
				Behaviour-Production				Pressure				Video camera	
	Kinetic	Kinematic	Indirect	Accelerometers	Milk metres	Milking robot	Feed scale	Weight scale	Pressure sensitive walkway	Force plate	Weight recording unit	Two dimensional (2D)	Three dimensional (3D)
Authors (year)													
Rajkondawar et al. (2002a)	x									x			
Pastell and Kujala (2007)	x										x		
Borderas et al. (2008)			x			x							
Song et al. (2008)		x										x	
Pastell et al. (2009)		x		x									
Chapinal et al. (2010)			x	x							x	x	
Ito et al. (2010)			x	x									
Poursaberi et al. (2010)		x										x	
Blackie et al. (2011a)			x	x	x								
Blackie et al. (2011b)		x		x								x	
Maertens et al. (2011)		x							x				
Pluk et al. (2012)		x							x			x	
de Mol et al. (2013)			x	x		x	x						
Kamphuis et al. (2013)			x	x	x			x					
Van Hertem et al. (2013b)			x	x	x								
Viazzi et al. (2013)		x										x	
Viazzi et al. (2014)		x											x

Table 2: Overview of the variables (gait, posture, behaviour-production and speed) used in the ALDM published in literature since 2002.

Model	Gait														Behaviour-Production													
	Ground reaction force	Stance time	Leg weight ratio	Number of kicks	Number of steps	Std.Dev. weight rear legs*	Tracking-up measurement	Variance acceleration	Stride length	Touch and release angle	Asymmetry step length	Asymmetry step time	Asymmetry step width	Asymmetry stance time	Back posture	Activity	Standing time	Lying time	Lying bouts number	Lying bouts time	Concentrate left over	Milking order	Milking frequency	Milk yield	Milking time	Live weight	Rumination	Speed
Authors (year)																												
Rajkondawar et al. (2002)	x	x																										
Pastell and Kujala (2007)			x	x	x	x																						
Borderas et al. (2008)																							x	x				
Song et al. (2008)							x																					
Pastell et al. (2009)								x																				
Chapinal et al. (2010)							x													x								x
Ito et al. (2010)																		x	x	x								
Poursaberi et al. (2010)															x													
Blackie et al. (2011)																	x	x	x	x				x				
Blackie et al. (2011)									x								x	x	x	x								x
Maertens et al. (2011)											x	x	x	x														
Pluk et al. (2012)										x																		
de Mol et al. (2013)																	x	x	x		x			x				
Kamphuis et al. (2013)																x						x		x	x	x		
Van Hertem et al. (2013)																x								x			x	
Viazzi et al. (2013)															x													
Viazzi et al. (2014)															x													

*Std.Dev. = standard deviation

Table 1 clearly indicates the different possibilities regarding ALDM. The kinetic approach was used first but since 2007 no new models based on this approach have been developed. The kinematic and the indirect approach on the other hand have been used alternately between 2008 and 2014. The most popular sensors are accelerometers and the use of video cameras is common as well. In the most recent ALDM, 2D recordings have been replaced by 3D recordings solving the problems due to a dynamic background.

Table 2 illustrates the difference between ALDM regarding the variables used to detect lameness. The use of activity or activity related variables is widespread. From the 17 ALDM listed in Table 1 and Table 2, seven models use this type of behavioural information. Three of the remaining models are based on measurements of one single variable: back posture. The potential of an ALDM combining back posture and activity related variables has not been investigated yet. This is remarkable since both variables have proven their use in the development of ALDM. By combining the two variables, new possibilities can be explored.

Walking speed measurements are only used in the models developed by Chapinal et al. (2010) and Blackie et al. (2011b). In the ALDM developed by these authors 2D, side-view video recordings were used to collect walking speed measurement and walking speed was only used as one of the variables. A profound investigation of the potential of walking speed in ALDM, especially based on 3D video recordings, has not been performed. Speed measurements however have been reported very promising for future lameness detection applications by Maertens et al. (2011).

3. Hypothesis and objectives

In this study three original feature variables were used in the development of automated lameness detection models (ALDM). Activity measurements (in steps per hour) and back posture measurements (dimensionless) were investigated and the performance of ALDM based on these variables was determined. The effect of the combination of two types of variables was investigated.

The most important innovative aspect of this study lies in the use of the third original feature variable: walking speed (in pixels per second) extracted from 3D video recordings. The performance of ALDM, involving this new variable, was also determined and compared with the performance of the ALDM based on back posture and activity measurements only.

The hypothesis investigated in this study is that walking speed is an interesting variable in the development of ALDM. The use of walking speed, extracted from 3D video recordings, as an additional original feature variable, is expected to increase the performance of the ALDM.

Four different objectives were set to gradually investigate the hypothesis:

1) Prove that walking speed of lame cows is significantly lower than walking speed of non-lame cows and that this difference in walking speed can be related to a difference in lameness state. Manual measurements were done to investigate this statement.
2) Demonstrate that useful walking speed measurements can be collected automatically.
3) Develop a reliable ALDM based on the common original feature variables back posture and activity with a specificity value of 80% and a sensitivity that is at least as good as the values obtained using manual systems (approximately 30% according to Espejo et al. (2006) and Fabian et al. (2014)).
4) A last objective is to prove that the use of walking speed can increase the specificity and the sensitivity of ALDM with at least 5% to 85% and 35% respectively.

4. Materials and methods

The main goal of this study was to investigate the potential of walking speed as a new variable in ALDM. In order to explore the walking speed variable and to investigate the effects of lameness state on walking speed, a specific experimental protocol was worked out. Data obtained from this experiment were used in the analysis of the first two objectives. During the experiment, manual and automated walking speed measurements were collected. The collected measurements were analysed independently. For the last two objectives, data were collected automatically over a four month period.

4.1 Animals and housing

All data used in this study were collected on a commercial dairy farm in Arendonk, Belgium. During the experiment and throughout the data collection period, the herd consisted of approximately 260 Holstein-Friesian cows divided in two groups based on their lactation stage. The animals involved in this study were randomly selected from both groups. The cows were housed in a cowshed with a total ground surface of 4000m² (100m x 40m). In the area reserved for the cows, with a surface of 2250m² (75m x 30m), slatted floors were used in the walking areas. Cows had access to cubicles with rubber mats and sawdust as bedding material. The sides of the shed could be opened or closed depending on the weather conditions.

A roughage mixture was fed to the cows once a day and pushed up twice every day. Next to roughage, a concentrate feed supplier was available to provide the amount of concentrate feed allocated to each individual cow. Two different mixes, with different protein contents, were used. Depending on the lactation stage and the milk yield, the amount and the composition of concentrate feed allocated to each cow were calculated automatically.

Milking took place twice a day in a rotary milking parlour with 40 places. The sessions were expected to start every day around 06.00h and 18.00h.

4.2 Experimental setup and execution of the walking speed experiment

To investigate the effect of lameness state on walking speed, an experimental protocol was developed. The preparations and the execution of this experiment are explained in this section.

4.2.1 Determination of sample size

The number of cows involved in the experiment that was needed to obtain statistical reliable results was determined first based on the estimation approach as described by Kutner et al. (2005). The following assumptions were made:

1) The power of the experiment was chosen at 90%. If the lameness state of a cow affects her walking speed, this effect will be detected with a 90% probability.
2) The significance level was chosen at 5% meaning that non-existing differences will be found in 5% of the cases.
3) The range in which walking speed and standard deviation on walking speed could be expected was extracted from previous studies:
    a. In the preparations of the experiment the expected difference in walking speed (notation Δ) between lame and non-lame cows was chosen at 0,30 meters per second (m/s). This value is intermediate to the values measured by Chapinal et al. (2010) and Blackie et al. (2011b) being 0,14m/s and 0,55m/s respectively.
    b. The standard deviation (SD) on walking speed was chosen at 0,15m/s based on previous research executed by Herlin and Drevemo (1994).
4) In this study the number of treatments (notation r) corresponds with the number of lameness states examined. A manual locomotion scoring session based on the system developed by Sprecher et al. (1997) was organised for the entire herd. The asymmetry of the gait, the stride length, the reluctance to bear weight and the curvature of the back were observed to assign a locomotion score (LS) between one and five to each cow. LS1 and LS5 were used for healthy and very lame cows respectively. Based on the LS assigned to the cows during this session, three different groups were distinguished: cows with LS1 and LS2 were considered non-lame, cows with LS3 were assigned to the moderately lame group and the third group consisted of severely lame cows with LS4 and LS5. An analysis comparing all five LS separately was not possible due to the low number of cows with LS4 and LS5 in the herd. A distinction between moderately and severely lame cows was made to guarantee that all degrees of lameness were represented in the experiment. For the actual analysis of the effect of lameness state (binary: non-lame versus lame) on walking speed, moderately and severely lame cows were combined in one group of lame cows.

Based on these four assumptions, the number of cows needed to perform a reliable experiment was extracted from the statistical table presented in Table 3 (Kutner et al., 2005).

Table 3: Determination of sample sizes for each group based on the estimation approach after Kutner et al. (2005). The black arrow indicates the situation used in the selection of the cows. The red arrow indicates the situation obtained after combining moderately and severely lame cows in one group.

	Power = 90%			
r	α = 0,20	α = 0,10	α = 0,05	α = 0,01
2	4	6	→ 7	10
3	5	7	→ 8	11
4	6	7	9	12
5	6	8	9	12

As indicated by the black arrow in Table 3, eight cows per group were required. Assuming equal sample sizes, a minimum of 24 cows, equally distributed over the three groups, was needed. With the help of the farmer, 24 plus eight additional cows were selected from the herd. The additional cows were selected to make sure that enough reliable measurements representing each of the lameness states would be retained. As indicated by the red arrow in Table 3, seven cows per group were needed to compare non-lame versus lame-cows.

4.2.2 Experimental setup for walking speed measurements

The trajectory used for the walking speed measurements was determined. To allow accurate measurements, the trajectory was chosen in the alley connecting the milking parlour with the housing pen. Walls at both sides of the alley forced the animals to complete the trajectory in a straight line optimising the measurements. The total length of the trajectory was equal to 16,43m. The position of the experimental setup is displayed in Figure 1.

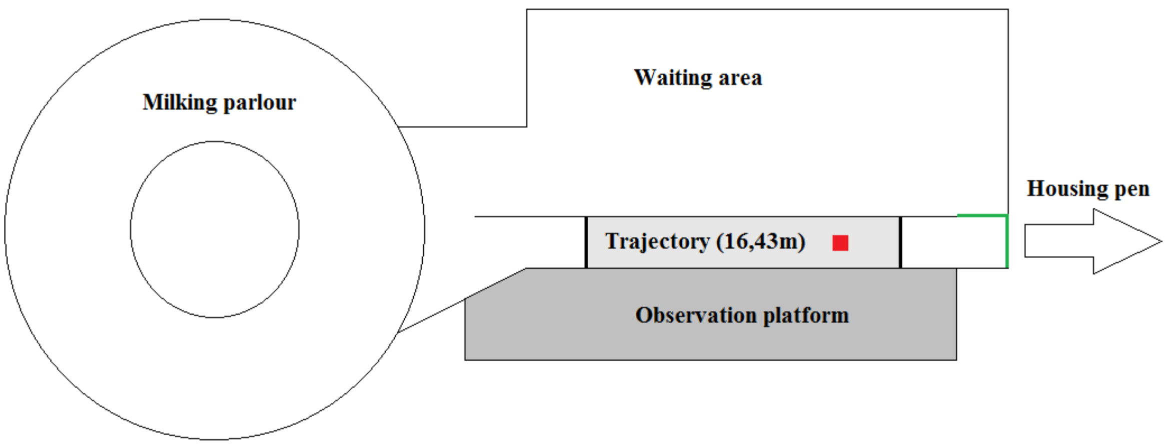

Figure 1 : Position of the trajectory in the shed. The light grey area represents the trajectory. The position of the video camera is indicated by the red square and the green lines show the position of the selection gates.

In order to investigate the second objective, regarding the automation of walking speed measurements, automated walking speed measurements were needed. Therefore a commercial video camera (Kinect^{TM}, Microsoft Corp., Redmond USA) was used. The camera was installed at a height of 3,20 meters above ground level to capture the entire body of the cows in the field of view of the camera. Projected at ground level, the field of view of the camera was equal to 3,74m. The position of the camera in relation to the trajectory is depicted by the red square in Figure 1.

4.2.3 Execution of the walking speed experiment

The experiment was conducted on July 16th 2014. All the cows (n = 32) selected for the experiment were separated from the rest of the herd and were diverted to the waiting area. In order to guide the cows through the trajectory one at a time and in a fluent way, three persons were needed. Each of them was responsible for a specific task:

1) Person 1 brought the cows to the entry of the alley one at a time.
2) Person 2 performed a manual locomotion scoring based on the system described by Sprecher et al. (1997). This scoring was used as a gold standard to identify the lameness states of the cows involved in the experiment. 15 cows were classified non-lame, eight cows were scored moderately lame and nine cows were identified severely lame. The minimal number of seven non-lame and seven lame cows required for analysis was thus fulfilled.
3) Once the animal arrived at the entry of the alley, person 3 followed the cow all the way through the trajectory. By doing so the cows were encouraged to walk the entire trajectory fluently. This third person was also responsible for the time measurements used for the calculation of the manual walking speeds: every time a cow passed the starting point, the digital chronometer (IKM882A, Calypso, Barcelona, Spain) was started and the time (in seconds) needed to reach the end of the trajectory was measured.

As soon as one cow completed the trajectory, the next cow was brought to the beginning of the alley and a new measurement was started. This procedure was repeated for all 32 cows consecutively. It took about 30 minutes to lead the 32 cows through the alley once. To increase the statistical reliability of the experiment, each cow was led through the alley three times. After the first round, the cows were led directly back into the waiting area by opening the selection gate (indicated by the green lines in Figure 3) connecting the alley and the waiting area. The same procedure was followed after the second round. After the third measurement another selection gate, connecting the alley with the housing pen, was opened and the animals returned to the herd. After 90 minutes, a total of 96 time measurements was collected.

Automated walking speed measurements were collected at the same time. Every time a cow completed the trajectory, she passed an electronic RFID antenna that triggered the recording of a video with the 3D video camera installed in top-view above the trajectory. The video recordings made during the experiment were processed and walking speeds were calculated using the commercial software package MATLAB version 8.2.0.701 (The MathWorks Inc., 2013). Automated walking speed measurements were expressed in pixels per second. Although 32 cows passed the camera three times each, only a total of 90 measurements for automated walking speed was retained after analysis. The remaining six recordings could not be used due to bad transponder connections (leading to missed identification of the animals) or inferior quality of the video images. For 27 cows (12 non-lame, eight moderately lame and seven severely lame) all three video recordings were useful. The required number of seven cows per group was met for the automated walking speed measurements.

4.3 Statistical analysis of experimental measurements

The measurements collected during the experiment described above were processed. Both the manual and the automated walking speed measurements were analysed statistically. The statistical analysis was done using the statistical R-software version 3.1.2 (R Development Core Team, 2011). Three different factors were distinguished: *'Lameness state'* representing the lameness state of the cows, *'Cow'* (nested in lameness state) accounting for the individual character of each animal and *'Repetition'* as an indicator for the three repetitions per cow.

4.3.1 Preparation of measurements

Manual walking speed was calculated as the length of the trajectory divided by the time needed to complete it. Each of the 96 time measurements gave an original manual walking speed measurement (units m/s). The 90 automated walking speed measurements extracted from the video recordings (pixels/s) were used as original automated walking speed measurements. Since a two factor ANOVA analysis with repeated measurements on one factor (ANOVA repeated measures analysis), with increased robustness for balanced data, was used for the statistical analysis, only cows with three correct measurements were retained in the analysis of the automated measurements. For the original manual and the original automated walking speed measurements, the mean value for non-lame and lame cows was calculated.

The original automated walking speed measurements were expressed in pixels/s. The physically relevant unit for walking speed however is m/s. The relationship between these units was explored using the original automated walking speed measurements and the corresponding original manual walking speed measurements. The accuracy of this relationship was limited (appendix 1) and therefore the original unit of pixels/s was retained.

4.3.2 Normalisation procedure

An ANOVA repeated measures analysis requires the data to originate from a normally distributed population. The distribution of the original manual and the original automated walking speed measurements was checked visually and tested formally by means of a Shapiro-Wilk test (Kutner et al., 2005). In both cases, the outcome of this test suggested the data originated from a not normally distributed population. A transformation was applied on the measurements to normalise the data. The transformations have to be seen in a mere statistical context. They are needed in order to satisfy the requirements of the ANOVA repeated measures analysis but they do not influence the outcome of the analysis as proven in appendix 2. The original manual walking speed measurements were transformed using Equation 1 and Equation 2 was used for the transformation of the original automated walking speed. The created variables, obtained after the transformations, are given in Table 4.

$$Normalised\ manual\ walking\ speed\ = \frac{1}{(original\ manual\ walking\ speed)^2} \qquad \text{Equation 1}$$

$$Normalised\ automated\ walking\ speed\ = \frac{1}{(original\ automated\ walking\ speed)} \qquad \text{Equation 2}$$

From Equation 1 and Equation 2, the units of the normalised walking speed measurements can be determined. The normalised manual walking speed (MWS) measurements are expressed in s^2/m^2 and the units of the normalised automated walking speed (AWS) measurements are s/pixel. The mean value for MWS and AWS was calculated for non-lame and lame cows.

Table 4 : Summary of the original variables obtained from the experiment, the transformations applied on these variables as well as the resulting normalised variables, their units and their abbreviations.

Original variable	Transformation	Normalised variable	Units	Abbreviation
original manual walking speed	Equation 1	normalised manual walking speed	s^2/m^2	MWS
original automated walking speed	Equation 2	normalised automated walking speed	s/pixel	AWS

4.3.3 Analysis of normalised manual measurements

The MWS measurements were analysed to investigate the first objective. The MWS measurements were checked visually using a boxplot and tested formally using an ANOVA repeated measures analysis. The effect of lameness state on walking speed was tested. The ANOVA repeated measures analysis was also used to determine the contribution of the factors *'Lameness state'*, '*Cow*' and '*Repetition*' as well as the interactions between these factors to the total variance. Using the mean sum of squares values calculated in the ANOVA repeated measures analysis, the total variance was divided into two parts: the between cows variance and the within cows variance. The bigger the contribution of differences between different cows (between cows variance), the more reliable the measurements. The flowchart of the statistical procedure for the analysis of the MWS measurements is presented graphically in Figure 2.

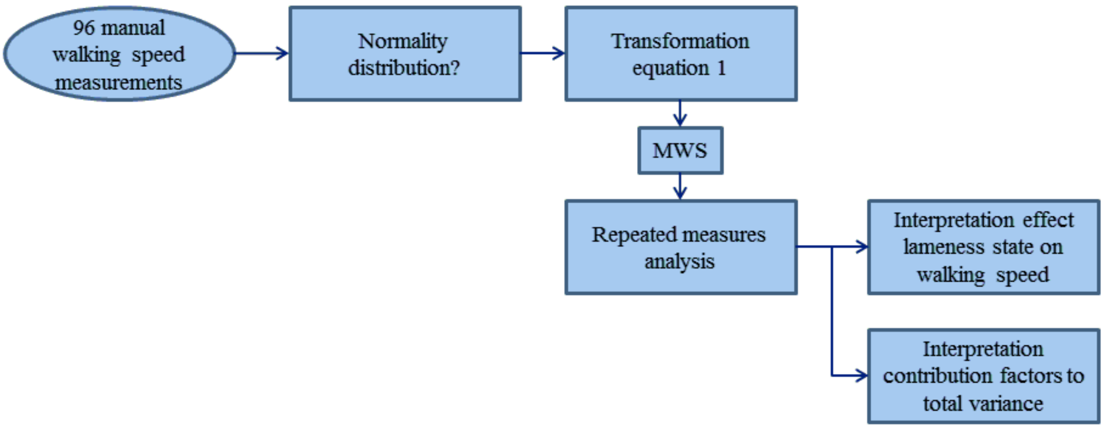

Figure 2 : Flowchart of the statistical procedure followed for the analysis of the manual walking speed measurements.

4.3.4 Analysis of normalised automated measurements

In order to use walking speed as a feature variable in the development of an ALDM, it has to be measured automatically. The collected AWS measurements were analysed to investigate the second objective. This analysis was run to see whether the same results as obtained from the MWS measurements could be extracted from the AWS measurements. The statistical procedure was identical as described above for the MWS measurements.

4.4 Preparations development of new ALDM

After investigating the relationship between walking speed measurements and lameness state based on the information collected during the experiment, the ALDM mentioned in the third and fourth objective were constructed using data collected between April 4th 2014 and July

31st 2014. All the cows in the herd were treated by a professional claw trimmer on April 2nd or April 3rd 2014. This treatment guaranteed a similar state for all the cows at the start of the data collection period. The end date was chosen two weeks after the walking speed experiment (July 16th 2014) to guarantee that the walking speed experiment was representative for the data collection period.

4.4.1 Collection data original feature variables

In this study the focus was on three original feature variables: activity, back posture and walking speed. For back posture (dimensionless) the range is between 0,20 and 0,65 and activity measurements vary between 0 and 250 steps per hour. Walking speed has a range between 0,1 and 0,75 pixels per second (around 0,60 and 2,00 m/s).

The same camera setup as used during the experiment provided video recordings of the cows twice a day, after the morning and evening milking session, when the cows passed through the alley to return to the housing pen. These recordings were processed and walking speed as well as back posture were calculated using the commercial software package MATLAB version 8.2.0.701 (The MathWorks Inc., 2013). Measurements for the third feature variable, the activity of the cows, were provided by the herd management software (ALPRO™, DeLaval AB, Tumba, Sweden) on an hourly base. The measurements were collected from the accompanying activity sensor which was attached to the collar of each cow and processed by the herd management software itself.

During the data collection period, six manual locomotion scoring sessions were organised for all the cows in the herd. The lameness state of the cows was determined using the system developed by Sprecher et al. (1997) and was used as a gold standard in the development of the ALDM described further. For the cows involved in the walking speed experiment, the locomotion scoring session performed during the experiment was also used.

4.4.2 Processing feature variables

Under practical conditions, not every video recording provided useful information regarding walking speed and back posture. Cows hindering each other while passing underneath the camera caused inaccurate and therefore useless recordings. Missed identifications and inferior quality of the video images also reduced the number of useful recordings. In order to obtain reliable information, the correct recordings from the last week were combined to calculate one average value for walking speed over the last week (WS, in pixels/s). Every day the period used to calculate this average value was updated. The same procedure was used for the back posture of the cows (BPM, dimensionless). Another feature variable was created by calculating the ratio of back posture over walking speed (BPM / WS, in s/pixel).

The hourly activity measurements were combined in one value representing the average activity (ACT, in steps per day) over the last week. By doing so all feature variables were set on the same frequency. Every day a new value was calculated using information collected in the last seven days. The final feature variables, as created based on this procedure, are listed in Table 5. The created variables were eventually used in the development of the ALDM.

Table 5 : Overview of the original and the final feature variables obtained after processing. The abbreviations, units and a short description of each variable are also given.

Original feature variable	Final feature variable	Abbreviation	Units	Description
Back posture	Average back posture	BPM	[-]	Average value of the correct measurements from the last week.
Walking speed	Average walking speed	WS	pixels/s	Average value of the correct measurements from the last week.
/	Average back posture / Average walking speed	BPM / WS	s/pixel	Ratio of BPM over WS
Activity	Average activity	ACT	steps/day	Average value of the hourly measurements from the last week.

4.4.3 Selection cows for calibration

For the calibration of the ALDM, the same 32 cows as involved in the experiment were selected first. Since three original feature variables were investigated, measurements for these three feature variables were needed at any time. For activity the critical period without information was set at three days, for walking speed and back posture the limit was set at seven days. Data collected from the other 20 cows were used in the calibration of the ALDM. Twelve cows had missing data for at least one feature variable during the data collection period (possible reasons include sensor malfunction and dry-off).

The cows used in the calibration of the ALDM represented all degrees of lameness: four cows were never scored lame during the data collection period, 11 cows were scored both lame and non-lame and the remaining five cows were scored lame at every manual scoring session between April 4th and July 31st.

4.5 Development new ALDM

For the development of an ALDM, the measurements were processed by three different modelling strategies: statistical modelling, autoregressive modelling and autoregressive modelling with exogenous variables. Statistical models, focussing on the level of the entire group, were developed first. Time series analyses of the final feature variables and the interactions between the final feature variables were investigated using dynamic autoregressive (dAR) models and dAR models with exogenous variables (dARX).

4.5.1 Construction of statistical models

Logistic regression was used in the development of probabilistic statistical models. The general structure of these models is presented in Equation 3 and Equation 4. BPM, ACT and WS were used as explanatory variables. In combination with a constant term, these variables were combined in an explanatory matrix (X). Based on the gold standard and the values in X, the optimal value of the model parameters (c) was determined and the lameness probability (P) was estimated at any point in time for each cow. The smaller P, the smaller the chance a cow was lame. The model parameters c were the same for all the cows. Therefore statistical models can be seen as group models.

$$ln \frac{P}{1- P} = X * c \qquad \text{Equation 3}$$

$$P = \frac{e^{X*c}}{1+ e^{X*c}} \qquad \text{Equation 4}$$

To investigate the potential of WS in ALDM, a distinction between models involving WS and models without WS was made. In a first step all possible statistical models based on the common final feature variables (BPM and ACT) were investigated. Three models were distinguished: model 1 based on ACT, model 2 using only BPM and model 3 based on ACT as well as BPM.

Using WS, another four statistical models were created: model 4 using WS as a single variable, model 5 combining ACT with WS, model 6 based on BPM as well as WS and model 7 combining ACT, BPM and WS. The fourth final feature variable BPM / WS was not used in the statistical models.

4.5.2 Construction of dynamic autoregressive models

Using statistical models, the individual dynamics of each cow were not brought into account. To incorporate these individual dynamics, dAR models were developed. For each cow an individual set of time-varying (a new calculation was made every day) parameters was calculated. dAR models are always based on one single feature variable. The general formula for a dAR model of order n is given in Equation 5.

$$y(k) = \frac{1}{1 + a_1 z^{-1} + a_2 z^{-2} + \cdots + a_n z^{-n}} e(k) \qquad \text{Equation 5}$$

In Equation 5 the model output and the error at time instant k are given by y(k) and e(k) respectively. The a values represent the model parameters and z represents the backward time shift operator.

The four final feature variables listed in Table 5 were used separately in a dAR model. As explained for the statistical models a distinction between models with and without WS was made. Two models did not involve WS (model 8 based on ACT and model 9 using BPM). Model 10, investigating WS and model 11 based on BPM / WS did use the WS information.

4.5.2.1 Determination of optimal model structure dAR models

Before the cow individual sets of parameters were calculated, the general dAR model structure and the size of the moving windows (MW) used for parameter estimation and prediction were determined. This was done in three steps :

1) The optimal model structure (the number of a parameters) was determined. The size of the estimation and prediction MW was temporarily set at ten days and five days respectively. The number of parameters was varied between one and ten and the optimal number of parameters was determined based on the Akaike information criterion (AIC) (Akaike, 1974). Every day a new calculation was made. Eventually, the number of parameters that was selected in most cases was chosen and implemented in the general structure.
2) With the optimal number of parameters identified, the size of the MW for estimation and prediction was optimised. The size of the estimation and prediction MW was varied between five and 15 days to investigate the impact of regularly recurring influences. The mean squared errors in the prediction MW were calculated for each combination. Based on a visual representation (as shown in Figure 10) of the mean squared errors, the optimal size of the estimation and the prediction MW was determined.
3) The individual and time-varying parameters were calculated given the number of a parameters and the size of the MW.

4.5.3 Construction of dAR models with exogenous variables

The last type of models developed in this study were dAR models with exogenous variables (dARX). Some variables were expected to influence other variables. In dARX models the interactions between different variables are considered next to the dynamics. The influencing variable was used as an input and the influenced variable as an output in the same model. The general formula for a dARX model of order n is presented in Equation 6.

$$y(k) = \frac{b_0 + b_1 z^{-1} + b_2 z^{-2} + \cdots + b_m z^{-m}}{1 + a_1 z^{-1} + a_2 z^{-2} + \cdots + a_n z^{-n}} u(k - \delta) + \frac{1}{1 + a_1 z^{-1} + a_2 z^{-2} + \cdots + a_n z^{-n}} e(k) \qquad \text{Equation 6}$$

In Equation 6, the model input, the model output and the error at time instant k are given by u(k), y(k) and e(k) respectively. The a and b values represent the model parameters and z represents the backward time shift operator. The δ symbol represents the time delay.

Only dARX models based on one input (mentioned first in this study) and one output variable were considered during this study. Based on this approach, three additional models were developed. Model 12 did not involve WS and investigated the relationship between BPM and ACT. Models 13 and 14 did involve WS. Model 13 searched for the relationship between BPM / WS and ACT. In model 14 the relationship between BPM and WS was investigated.

4.5.3.1 Determination of optimal model structure dARX models

The same three steps as used for the determination of the structure of dAR models were followed to identify the optimal structure for the dARX models developed in this study. The only difference was found in step 1. Next to the optimal number of a parameters, the number of b parameters and the time delay δ were also optimised in this step for dARX models.

4.6 Performance assessment of the developed models

To compare the performance of the developed models, a receiver operating characteristic (ROC) curve (Kutner et al., 2005) was drawn and the Area Under the Curve (AUC) was calculated for each model. Based on the original gold standard, rough ROC curves were obtained. An additional assumption was made: if a cow was assigned the same LS in two consecutive scoring sessions, she was assumed to have the same LS during the entire period between the two sessions. This assumption was used for all the ROC curves presented in this study. For the calibration and calculation of model parameters, the original gold standard was used.

For the statistical models the calculated lameness probability scores were compared with the gold standard to draw the ROC curves. For the dAR and the dARX models the individual and time-varying parameters were used. The influence of a series of mathematical operations (presented in Table 6 and Table 7 for the dAR and the dARX models respectively) on the calculated parameter sets was also investigated. The same optimal model structure was found for all four the dAR models allowing the use of the same mathematical operations and combinations on each dAR model. The same was valid for the dARX models.

Table 6 : Overview of the cow individual parameters and the mathematical operations executed on these parameters for the dAR models. An abbreviation and a short explanation are given for each operation.

Mathematical operation	Abbreviation	Explanation
/	a_1	Course of original a_1 parameters
First derivative a_1	$d(a_1)$	Rate at which a_1 parameters change
Variance a_1 over the last week	$var(a_1)$	Measure for the irregularity of the a_1 parameters
Sum of original a_1 parameter and first derivative a_1	$a_1 + d(a_1)$	Combine information original a_1 parameters and rate of change
Model prediction error	Error	Investigate accuracy of model predictions

Table 7: Overview of the cow individual parameters and the mathematical operations executed on these parameters for the dARX models. An abbreviation and a short explanation are given for each operation.

Mathematical operation	Abbreviation	Explanation
/	a_1	Course original a_1 parameters
/	a_2	Course original a_2 parameters
/	b_0	Course original b_0 parameters
First derivative a_1	$d(a_1)$	Rate at which a_1 parameters change
First derivative a_2	$d(a_2)$	Rate at which a_2 parameters change
First derivative b_0	$d(b_0)$	Rate at which b_0 parameters change
Variance a_1 over the last week	$var(a_1)$	Measure for the irregularity of the a_1 parameters
Variance a_2 over the last week	$var(a_2)$	Measure for the irregularity of the a_2 parameters
Variance b_0 over the last week	$var(b_0)$	Measure for the irregularity of the b_0 parameters
Sum of a_1, a_2 and b_0 parameters	$a_1 + a_2 + b_0$	Combine the information in the different parameters
Sum of a_1 and a_2 parameters	$a_1 + b_0$	Combine information in parameters of highest order
Difference of $(a_1 + a_2)$ and b_0 parameters	$a_1 + a_2 - b_0$	Combine information in different parameters, counteracting
Model prediction error	Error	Investigate accuracy of model predictions

4.6.1 Selection of best models

14 different ALDM were developed in this study. The ALDM are listed and discussed per model type in the results section. An overview of the different models is given in appendix 3. Based on the ROC curves and the AUC values, the models were compared. Models with high AUC values (around 70%) were selected as promising models. In case of similar AUC values,

the easiest model (with the lowest number of feature variables or based on the easiest mathematical operation in case of dAR and dARx models) was preferred. Six models, the most promising model with and without WS for each model type, were selected and analysed.

For each of the selected models, the thresholds (used for classification of the cows) corresponding with a specificity of 90% were determined. The value of 90% creates a buffer for the 80% specificity after validation that was pursued according to the third objective. The thresholds corresponding with a sensitivity of 90% were also determined to investigate the performance of the models regarding the detection of lame cows.

4.7 Validation procedure

After calibration and selection of the most promising models, these models were validated. The validation procedure was run to compare the final performance of the different models and to identify the model with the highest performance.

4.7.1 Selection of validation dataset

20 new cows were used for validation of the selected models. These cows were randomly selected from the remaining cows in the herd. This was done in three steps:

1) Removal of the 32 cows involved in the walking speed experiment. The 20 cows used for model calibration belonged to this group of 32 cows and were removed at the same time. The remaining cows at this point had not been used in any way yet.
2) Removal of cows with missing data for at least one of the feature variables between April 4th and July 31st. The maximal period without information that was allowed was set at three days for activity and seven days for back posture and walking speed.
3) Random selection of 20 cows from the cows remaining after the first two steps.

The 20 cows used for validation did represent all different lameness states: based on the locomotion scoring sessions that were used as a gold standard, seven cows were never lame during the data collection period. Ten cows were scored lame as well as non-lame and the remaining three cows were always scored lame.

4.7.2 Validation of statistical models

The selected statistical models, both with and without WS, were validated in five steps:

1) The values of the c parameters in Equation 3 and Equation 4 were calculated based on the calibration data.

2) The c parameters obtained from step 1 were used to calculate the lameness probability scores of the cows selected for validation.
3) The lameness probability threshold (P_{thresh}) corresponding with a specificity of 90% (high specificity threshold) for the calibration data was used to classify the lameness probability scores (P) of the cows in the validation dataset. Cows with $P > P_{thresh}$ were classified lame, cows with $P < P_{thresh}$ were classified non-lame.
4) The model classifications were compared with the gold standard for the cows selected for validation. The specificity and the sensitivity were determined.
5) Step 3 and step 4 were repeated with the thresholds corresponding with a sensitivity of 90% (high sensitivity threshold). This step was used to analyse the performance of the models regarding the detection of lame animals.

4.7.3 Validation of dAR and dARX models

The promising dAR and dARX models were validated using the same procedure. This procedure consisted of five different steps. As shown in the results section (Table 13 and Table 14), the mathematical operations did not improve the AUC values. Therefore the thresholds mentioned in step 3 were applied on the original cow individual parameters.

1) The optimal model structure as determined for the calibration data was set and the cow individual parameter sets were calculated for the cows in the validation dataset.
2) Based on the AUC values calculated for the calibration data, the b_0 parameter was selected for validation.
3) The threshold on the b_0 parameters (B_{thresh}) corresponding with a specificity of 90% (high specificity threshold) for the calibration data was used to classify the b_0 parameters of the cows in the validation dataset. If the value of the b_0 parameter was higher than B_{thresh}, the cow was classified lame.
4) The model classifications for the cows in the validation dataset were compared with their gold standard and the specificity and sensitivity were calculated.
5) As explained for the statistical models, step 3 and step 4 were repeated with the thresholds corresponding with a sensitivity of 90% (high sensitivity threshold).

4.8 Interpretation of different model performances

The performance of the selected models was analysed. For each of the model types, the influence of WS on the model performance was determined. The calculated specificity and sensitivity values were compared with the results of the ALDM published in literature.

5. Results

The results of the different analysis steps in the construction of the ALDM developed in this study and the performance of these models are presented in this section.

5.1 Results effect lameness state on walking speed

The average value for the original manual walking speed measurements was equal to 1,29 m/s for non-lame cows. For lame cows an average value of 1,05 m/s was found. The difference between the original manual walking speed of non-lame and lame cows was 0,24 m/s. This value approached the value that was used (0,30 m/s) in the estimation approach to determine the statistical sample size.

5.1.1 Normalisation procedure

The distribution of the manual walking speed measurements is presented in Figure 3. Visual observation of the left part of Figure 3 shows that the deviation of the original manual walking speed measurements from the red line (indicating a perfectly normal distribution) is large. The normalised MWS measurements are shown in the right part of Figure 3.

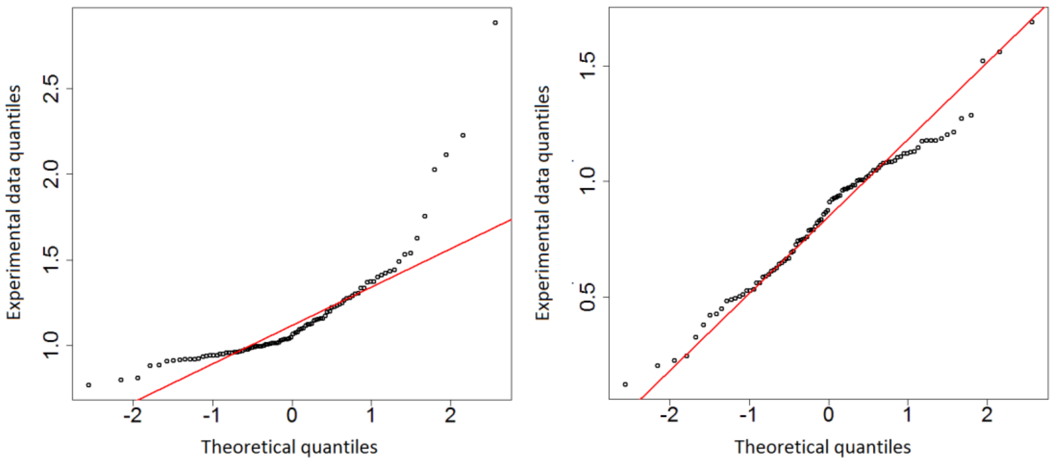

Figure 3 : Graphical representation of the distribution of the manual walking speed measurements. On the left the original measurements are given, on the right the normalised MWS measurements are presented. The red line represents a perfectly normal distribution.

The results of the formal Shapiro-Wilk test for the original and the normalised manual walking speed measurements are given in Table 8.

Table 8 : Overview of the results of the Shapiro-Wilk test investigating the distribution of the manual walking speed measurements.

	Test statistic W	P-value
Original manual walking speed	0,752	$1{,}898 * 10^{-11}$
Normalised manual walking speed (MWS)	0,983	0,237

The calculated P-value for the original measurements was smaller than the significance level of $\alpha = 0{,}05$. The null hypothesis, assuming the data originated from a normal distribution, was rejected. After normalisation the P-value for the MWS measurements was equal to 0,237. This value exceeded the significance level and the null hypothesis was accepted.

5.1.2 Effect of lameness state on MWS

The mean values for MWS were equal to 0,60 s^2/m^2 for non-lame cows and 0,91 s^2/m^2 for lame cows respectively. The effect of lameness state on MWS is presented as a boxplot in Figure 4.

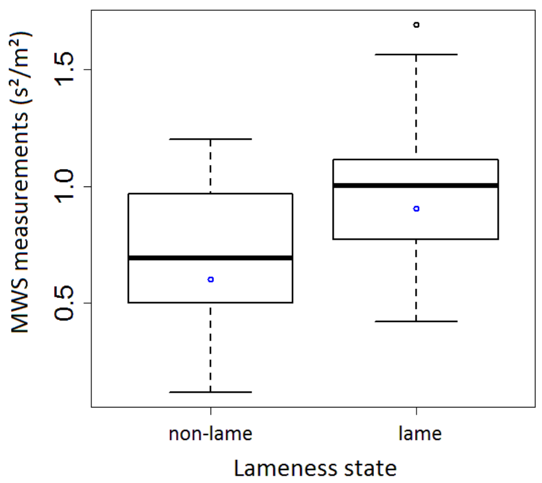

Figure 4 : Boxplot depicting the MWS measurements for lame and non-lame cows. The horizontal black line represents the median and the dot in the box gives the mean value for each group.

In Figure 4 the mean MWS of lame cows is higher than the mean MWS of non-lame cows. This is probably due to the use of the transformation presented in Equation 1. The MWS measurements were investigated formally using an ANOVA repeated measures analysis. The results of this test are given in Table 9.

Table 9: Results of the ANOVA repeated measures analysis for the MWS measurements. The factors are grouped per level and the number of degrees of freedom (df) as well as the sum of squares and the mean sum of squares are given for each factor. The P-values are also presented. Significant P-values are indicated with an asterisk.

Two factor ANOVA with repeated measurements on one factor						
Factor level		**df**	**Sum of squares**	**Mean sum of squares**	**F-test**	**P-value**
Between cows	**Lameness state**	1	1,475	1,475	7,604	0,010 *
	Cow (nested in lameness state)	30	5,820	0,194		
Within cows	**Repetition**	2	0,041	0,021	1,035	0,362
	Lameness state:Repetition	2	0,121	0,060	3,018	0,056
	Residuals	60	1,202	0,020		

The effect of the *'Lameness state'* factor on the MWS measurements was significant (P-value = 0,010). Other effects were not significant.

5.1.3 Reliability of MWS measurements

The three MWS measurements of the 32 cows involved in the walking speed experiment are presented in Figure 5 and allow a visual investigation of the reliability of the MWS measurements.

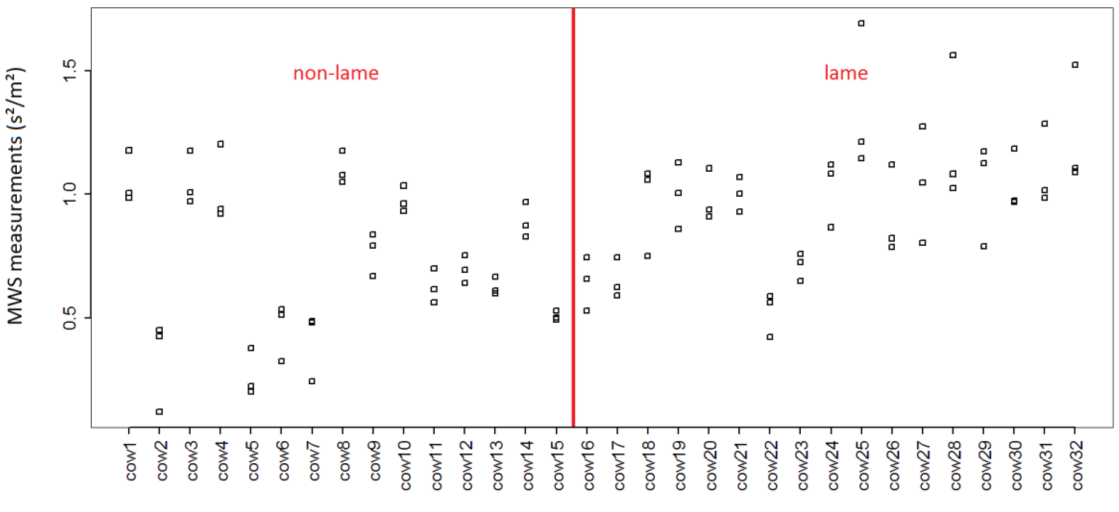

Figure 5 : Visual representation of the variance between and within cows for the MWS measurements. The non-lame cows are presented first (cow1 to cow15), followed by the lame cows (cow16 to cow32). The red line separates these two groups.

Visual observation suggests that the variance between different cows is bigger than the variance within one cow. An exact value for the proportion of the total variance declared by differences between and within cows was calculated using the mean sum of squares values given in Table 9. This is shown in Equation 7 and Equation 8.

$$\text{Proportion between cows variance} = \frac{1{,}475 + 0{,}194}{1{,}475 + 0{,}194 + 0{,}021 + 0{,}060 + 0{,}020} = 0{,}943 \quad \text{Equation 7}$$

$$\text{Proportion within cows variance} = \frac{0{,}021 + 0{,}060 + 0{,}020}{1{,}475 + 0{,}194 + 0{,}021 + 0{,}060 + 0{,}020} = 0{,}057 \quad \text{Equation 8}$$

The proportion of variance declared by differences between cows was equal to 94.3%. The proportion of variance declared by differences within cows was limited to 5,7%. The MWS measurements and the results obtained from the ANOVA repeated measures analysis of the MWS measurements were reliable.

5.2 Results collection automated walking speed measurements

For the automated measurements, the average value for the original walking speed was equal to 0,38 pixels/s for non-lame cows. The average value found for lame cows was lower (0,31 pixels/s). The difference between the two groups was equal to 0,07 pixels/s.

5.2.1 Normalisation procedure

The distribution of the automated walking speed measurements is presented in Figure 6. Visual observation of the left part of the graph shows that the deviation of the original walking speed measurements from the red line (indicating a perfect normal distribution) is large. The AWS measurements obtained after normalisation are displayed in the right half of Figure 6.

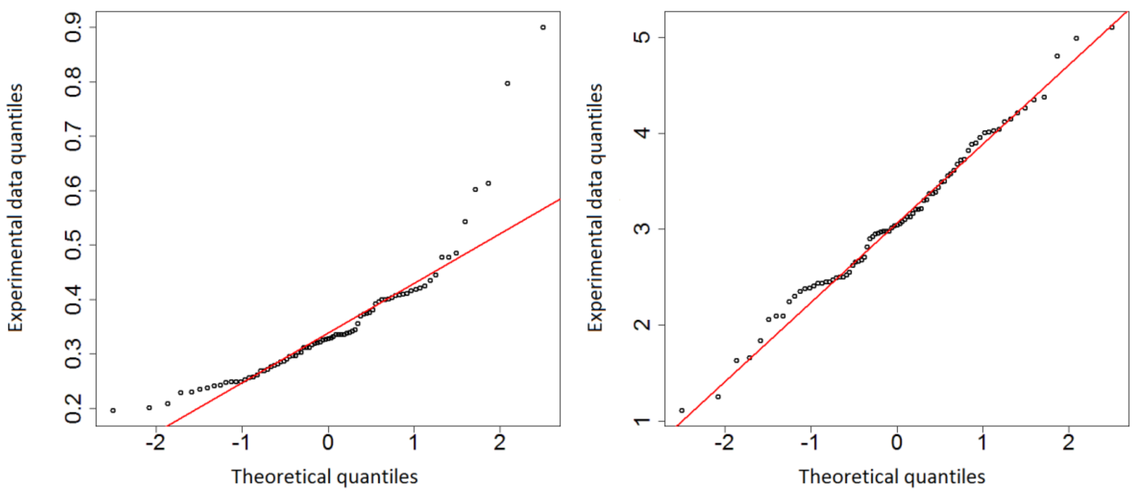

Figure 6 : Graphical representation of the distribution of the automated walking speed measurements. On the left the original measurements are given, on the right the normalised AWS measurements are presented. The red line represents a perfectly normal distribution.

The formal results regarding the distribution of the automated walking speed measurements are presented in Table 10. The results for the original walking speed measurements and the AWS measurements obtained after normalisation are given.

Table 10 : Overview of the results of the Shapiro-Wilk test investigating the distribution of the automated walking speed measurements.

	Test statistic W	P-value
Original automated walking speed	0,812	$9{,}007 * 10^{-9}$
Normalised automated walking speed (AWS)	0,990	0,777

For the original measurements the P-value was smaller than the significance level of $\alpha = 0{,}05$. The data did not originate from a normally distributed population. The P-value of the normalised AWS measurements was equal to 0,777. This value was bigger than the significance level confirming the AWS measurements were normally distributed.

5.2.2 Effect of lameness state on AWS

The mean values of the AWS measurements were equal to 2,58 and 3,18 s/pixels for non-lame and lame cows respectively. The relationship between lameness state and AWS measurements is presented in the boxplot shown in Figure 7.

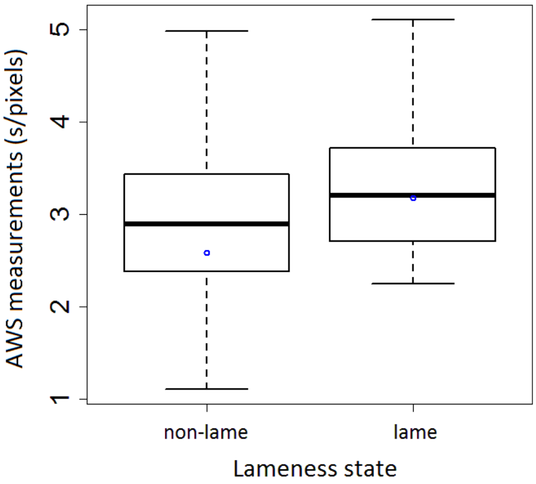

Figure 7 : Boxplot depicting the AWS measurements for lame and non-lame cows. The horizontal black line represents the median and the dot in the box gives the mean value for each group.

The mean AWS of lame cows lies lower than for non-lame cows (the transformation presented in Equation 2 has to be noticed). The difference between the mean values of both groups is small. The range of the AWS measurements is bigger for the non-lame cows. The results of the formal ANOVA repeated measures analysis for the AWS measurements are given in Table 11.

Table 11 : Results of the ANOVA repeated measures analysis for the AWS measurements. The factors are grouped per level and the number of degrees of freedom (df) as well as the sum of squares and the mean sum of squares are given for each factor. The P-values are also presented.

Two factor ANOVA with repeated measurements on one factor						
		df	**Sum of squares**	**Mean sum of squares**	**F-test**	**P-value**
Between cows	**Lameness state**	1	4,026	4,026	3,686	0,066
	Cow (nested in lameness state)	25	22,307	1,092		
Within cows	**Repetition**	2	1,359	0,680	1,814	0,174
	Lameness state:Repetition	2	0,266	0,133	0,354	0,703
	Residuals	50	18,734	0,375		

None of the factors listed in Table 11 was significant at the significance level used in this study. The smallest P-value was found for '*Lameness state'* (P = 0,066).

5.2.3 Reliability of AWS measurements

The AWS measurements of the cows with three correct repetitions are presented in Figure 8. The variance within the measurements of one cow and the variance between measurements of different cows were compared.

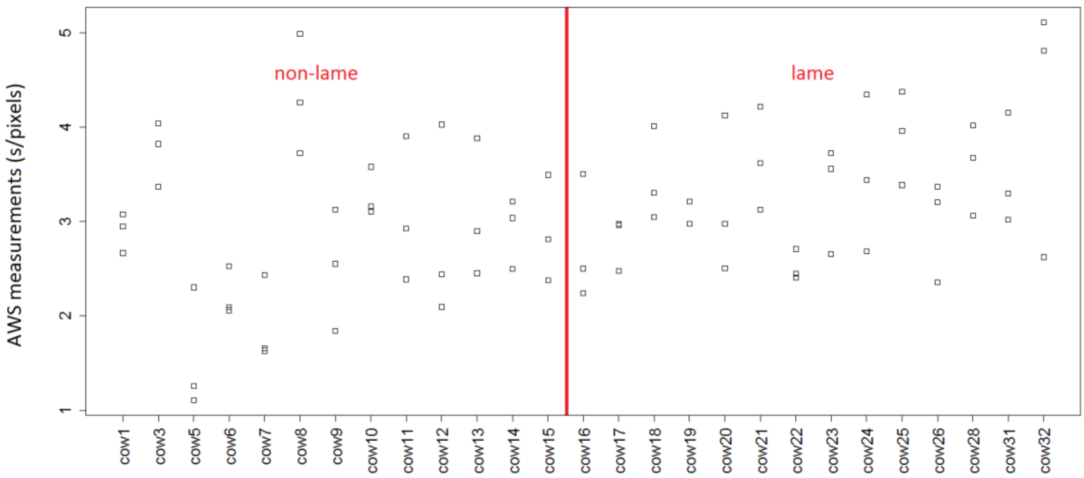

Figure 8 : Visual representation of the variance between and within cows for the AWS measurements. The non-lame cows are presented first (cow1 to cow15), followed by the lame cows (cow16 to cow32). The red line separates these two groups.

The variance within cows is bigger if compared to the MWS measurements (Figure 5). The exact proportions of between cows and within cows variance were calculated based on the mean sum of squares values given in Table 11. The calculations are given in Equation 9 and Equation 10.

$$Proportion\ between\ cows\ variance\ = \frac{4{,}026 + 1{,}092}{4{,}026 + 1{,}092 + 0{,}680 + 0{,}133 + 0{,}375} = 0{,}812 \qquad \text{Equation 9}$$

$$Proportion\ within\ cows\ variance\ = \frac{0{,}680 + 0{,}133 + 0{,}375}{4{,}026 + 1{,}092 + 0{,}680 + 0{,}133 + 0{,}375} = 0{,}188 \qquad \text{Equation 10}$$

The proportion of between cows variance dropped from 94,3% (for MWS measurements) to 81,2% for AWS measurements. The proportion of within cows variance increased from 5,7% to 18,8%. This value was still acceptable but the accuracy and the reliability of the AWS measurements was lower than for the MWS measurements.

5.3 Results development new ALDM

5.3.1 Results statistical models

The AUC values for the seven statistical models developed in this study are given in Table 12. The model parameters c are also given in Table 12. The first c parameter corresponds with the constant term, the other parameters correspond with the explanatory variables given in column 2 of Table 12. For model 2 for example this means that the second c parameter corresponds with BPM as illustrated in Equation 11.

Table 12 : Statistical models developed in this study. The explanatory variables as well as the calculated model parameters (c) and the AUC values are given for each model. The AUC values shown in bold indicate the models that were selected as most promising models.

Model	Explanatory variables	c parameters	AUC (%)
Model 1	ACT	[-0,37 0,01]	52,0
Model 2	BPM	[-4,89 14,38]	**68,6**
Model 3	ACT + BPM	[-5,89 0,01 14,88]	68,6
Model 4	WS	[3,33 -15,92]	67,1
Model 5	ACT + WS	[2,79 0,02 -19,50]	69,4
Model 6	BPM + WS	[-1,71 11,54 -11,22]	**71,2**
Model 7	ACT + BPM + WS	[-2,28 0,02 11,58 -14,85]	72,0

For the statistical models without WS, model 2 with an AUC value of 68,6% was selected as the most promising model. Model 3 had the same AUC value but an additional variable was required to obtain this result. Model 6 was selected as most promising model with WS involved. The AUC value of model 6 was equal to 71,2%. This value was 0,8% lower than for model 7 but ACT was not used in model 6. The AUC value of model 1 (52,0%), based on ACT only, was almost 20% lower than the AUC value found for model 6. The model structure of the selected models is presented in Equation 11 and Equation 12 for model 2 and model 6 respectively. In Equation 12 the sign of the coefficient in front of BPM is positive and WS has a negative coefficient suggesting that BPM and WS are affected in opposite ways. High lameness probabilities are found if BPM is high and WS is low.

$$ln\frac{P}{1-P} = -4{,}89 + 14{,}38 * BPM \qquad \text{Equation 11}$$

$$ln\frac{P}{1-P} = -1{,}71 + 11{,}54 * BPM - 11{,}22 * WS \qquad \text{Equation 12}$$

The ROC curves of model 2 and model 6 are depicted in Figure 9. The blue curve is the curve of model 2 and the red curve represents model 6. In general the red curve lies higher than the blue curve hence the higher AUC-value.

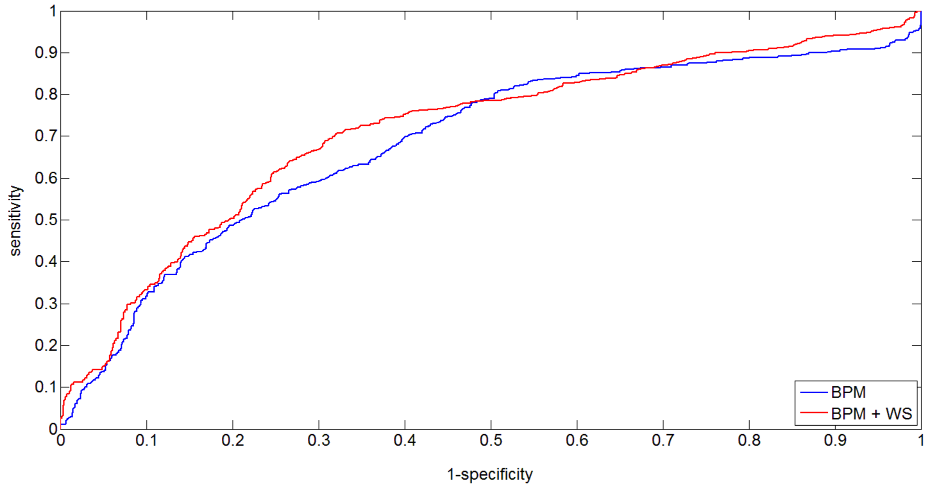

Figure 9 : ROC curves of the selected statistical models. The blue line shows the curve of model 2 and the curve for model 6 is shown in red. On the vertical axis the sensitivity is given. The horizontal axis contains 1-specificity values. The variables used in the models are shown in the legend.

5.3.2 Results dAR models

5.3.2.1 Optimal model structure dAR models

The identified optimal model structure was identical for the four dAR models developed in this study. In 100% of the cases, a structure using one a parameter (a_1) was determined as the optimal structure for the dAR models. The general structure of a dAR model of order n as given in Equation 5 was simplified to a first order formula as presented in Equation 13.

$$y(k) = \frac{1}{1 + a_1 z^{-1}} e(k)$$ Equation 13

The optimal size of the moving windows (MW) for estimation and prediction of the parameters was determined from a visual representation of the mean squared errors. For model 9, based on BPM, the resulting plot is shown in Figure 10. The longer the estimation MW, the smaller the error. For smaller estimation MW the impact of an additional day is large. The influence of the prediction MW is less clear. Similar figures were found for the other dAR models. The optimal size of the MW for estimation was determined at twelve days. The prediction MW was set at seven days. This choice gives the farmer a few days to schedule the treatment in his routine but still guarantees an early treatment for the animal. These values were used for all the dAR models.

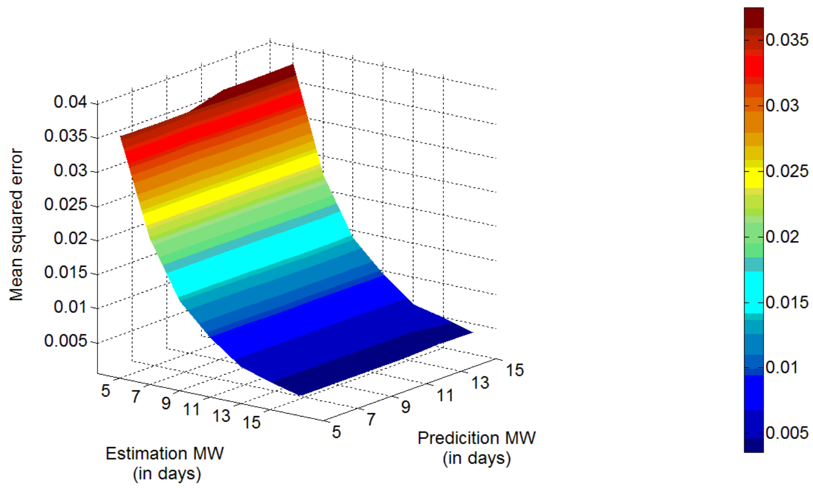

Figure 10 : Mean squared errors of the dAR model based on BPM in function of the size of the MW used for estimation and prediction of the parameters.

5.3.2.2 Selection dAR models

The AUC values calculated for the dAR models are listed in Table 13.

Table 13 : AUC values for the resulting parameters obtained from the mathematical operations. The results are given for all four the dAR models. The variables used in each model are presented in the second column. The values shown in bold indicate the combinations (model and mathematical operation) that were selected.

		AUC (%) values for the mathematical operations tested for dAR models				
	Used variables	a_1	$d(a_1)$	$var(a_1)$	$a_1 + d(a_1)$	Error
Model 8	ACT	50,7	51,5	52,2	50,1	50,7
Model 9	BPM	**55,6**	51,4	51,6	54,4	52,7
Model 10	WS	54,7	51,4	50,1	54,0	51,9
Model 11	BPM / WS	**56,6**	52,1	50,1	54,6	55,3

The original a_1 parameters gave the best result for three models. The mathematical operations did not improve the AUC values. For the dAR models without WS, the highest AUC value was found for model 9 (55,6%). The best result was found for model 11 (AUC of 56,6%). The AUC values presented in Table 13 are smaller than the AUC values of the statistical models given in Table 12. The use of dAR models did not increase the performance compared to the statistical models. The ROC curves of the best dAR models are given in Figure 11.

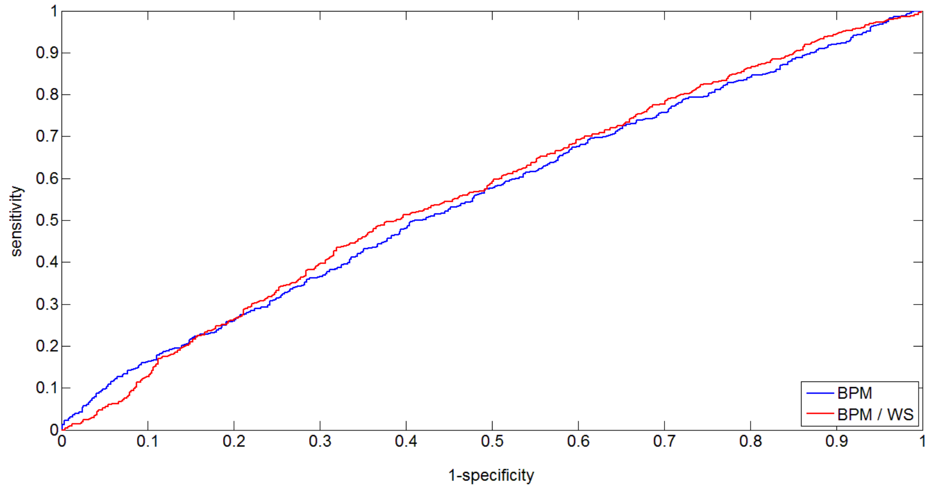

Figure 11 : ROC curves of the selected dAR models. The blue line shows the curve of model 9 and the curve of model 11 is shown in red. On the vertical axis the sensitivity is given. The horizontal axis contains 1-specificity values. The variables used in the models are shown in the legend.

5.3.3 Results dARX models

5.3.3.1 Optimal model structure dARX models

For the three dARX models the same optimal model structure was found: two a parameters (a_1 and a_2), one b parameter (b_0) and a time delay equal to zero. The general structure presented in Equation 6 was simplified to Equation 14.

$$y(k) = \frac{b_0}{1 + a_1 z^{-1} + a_2 z^{-2}} u(k) + \frac{1}{1 + a_1 z^{-1} + a_2 z^{-2}} e(k) \qquad \text{Equation 14}$$

The optimal size of the MW for estimation and prediction of the parameters was determined from a graphical representation similar to the one in Figure 10. For dARX models the optimal size of the MW for estimation was determined at 12 days. As explained for the dAR models, the prediction MW was set at seven days. The same values were found for the dAR models.

5.3.3.2 Selection dARX models

The AUC values calculated for the dARX model are listed in Table 14. The results for all the mathematical operations are given for the three dARX models. The variables that were used as input and output are also shown.

Table 14 : AUC values for the dARX models developed in this study. The results are given for the three dARX models. The variables used in each model are presented in the second column. The values shown in bold indicate the combinations (model and mathematical operation) that were eventually selected.

		Model 12		Model 13		Model 14	
Used variables	**Mathematical operation**	input	BPM	input	BPM / WS	input	BPM
		output	ACT	output	ACT	output	WS
AUC (%) values for the mathematical operations tested for dARX models	a_1	50,9		51,4		55,4	
	a_2	50,7		50,6		50,3	
	b_0	**58,0**		55,7		**67,2**	
	$d(a_1)$	50,8		51,3		50,7	
	$d(a_2)$	50,5		50,3		51,5	
	$d(b_0)$	51,0		50,3		50,4	
	$var(a_1)$	52,2		50,2		52,9	
	$var(a_2)$	53,5		53,1		50,4	
	$var(b_0)$	51,4		53,6		52,7	
	$a_1 + a_2 + b_0$	58,0		55,6		67,5	
	$a_1 + b_0$	58,0		55,5		68,1	
	$a_1 + a_2 - b_0$	58,0		55,8		63,0	
	Error	53,3		52,7		67,2	

Model 12 was the best dARX model without WS. The b_0 parameter gave the best result with an AUC value of 58,0%. The mathematical operations did not lead to higher AUC values. The best dARX model was model 14. WS was used in this model and AUC values between 67% and 68% were achieved. The AUC value obtained from the b_0 parameter (67,2%) was selected. Other AUC values were up to 0,9% higher but the mathematical operations were more complex. Compared to model 12, the AUC value calculated from the b_0 parameter was more than 9% higher for model 14. The AUC value of 67,2% for model 14 was 10,6% higher than the highest AUC value found for dAR models. The calculated values are similar to the AUC values obtained with statistical models. The ROC curves of model 12 and model 14 are shown in Figure 12. The curve of model 14 lies higher than the curve of model 12 at any point. The difference between the selected model with and without WS is bigger than for statistical and dAR models.

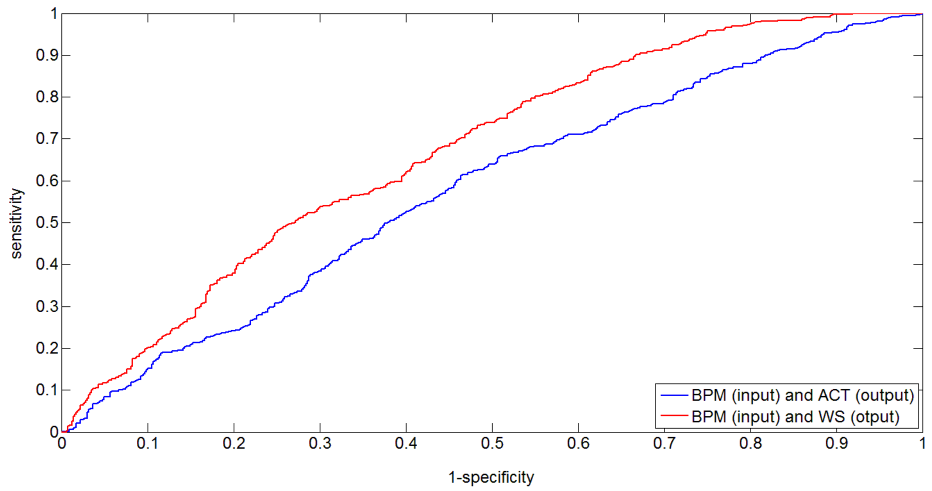

Figure 12 : ROC curves of the selected dARX models. The blue line shows the curve of model 12 and the curve of model 14 is shown in red. On the vertical axis the sensitivity is given. The horizontal axis contains 1-specificity values. The variables used in the models are shown in the legend.

5.3.4 Validation of selected models

5.3.4.1 Validation model specificity

The models were validated first using high specificity thresholds. The resulting specificities are listed in column five of Table 15. The sensitivities corresponding with these high specificity values are also presented in column six of Table 15. For the statistical models specificity values around 80% were reached. Without WS the specificity was equal to 79,4%. The use of WS led to a 2% increase in specificity from 79,4% to 81,6%. The corresponding sensitivity was low with values around 20%.

The influence of WS in dAR models was less clear. In Figure 11 the difference between the two ROC curves was small. This was confirmed by the values shown in Table 15. For model 9 specificity and sensitivity were equal to 72,1% and 21,4% respectively. For model 11 similar values of 71,8% and 23,0% were found. The specificity values calculated for the dAR models were worse than those calculated for the statistical models. For the sensitivities the results of the dAR and the statistical models were similar. In dARX models the use of WS led to an increase of 3% in specificity: from 76,4% to 79,5%. The sensitivities of model 12 and model 14 were equal to 24,1% and 22,3% respectively. The results found for the dARX models were better than the results calculated for the dAR models and approached the results found for the statistical models.

Regarding the specificity, the best ALDM developed in this study was model 6 with a specificity of 81,6%. The sensitivity of this model was equal to 22,6%. Similar values were found for model 14 but model 6 was easier and therefore preferred. The required computational power was also lower for model 6.

Table 15 : Overview of the specificity and the sensitivity of the six selected models. For each model type the results of the best model with and without WS are presented. Two possibilities are given: high specificities with the corresponding sensitivity values are given in the fifth and sixth column. In column 8 high sensitivities were targeted. The corresponding specificity values are given in column 7.

Model information			**Dataset**	**Results high specificity thresholds**		**Results high sensitivity thresholds**	
	Model name	Used feature variable(s)	Calibration (C) or Validation (V)	Specificity (%)	Sensitivity (%)	Specificity (%)	Sensitivity (%)
Statistical model	Model 2	BPM	C	90,9	20,8	18,2	90,7
			V	79,4	22,6	46,6	77,1
	Model 6	BPM + WS	C	90,3	16,0	9,1	90,7
			V	81,6	18,3	26,7	82,9
dAR model	Model 9	BPM	C	90,4	14,8	9,4	90,5
			V	72,1	21,4	4,0	73,3
	Model 11	BPM / WS	C	90,3	18.7	14,8	90,6
			V	71,8	23,0	20,0	73,3
dARX model	Model 12	BPM (input)	C	90,6	23,6	18,5	90,6
		ACT (output)	V	76,4	24,1	52,0	73,3
	Model 14	BPM (input)	C	90,4	23,1	22,2	90,5
		WS (output)	V	79,5	22,3	20,0	80,0

5.3.4.2 Validation model sensitivity

The performance of the automated lameness detection models developed in this study to detect lame cows was investigated from the sensitivities of the selected models. These results are listed in the last two columns of Table 15. For model 2 a sensitivity of 77,1% was reached. In model 6 the use of WS led to an increase in sensitivity of almost 6%: from 77,1% to 82,9%. The specificity dropped by almost 20% from 46,6% to 26,7%.

For dAR models the use of WS did not change the sensitivity of the models. For both models the sensitivity was equal to 73,3%. This value was 9,6% lower than the sensitivity of model 6. The specificity increased from 4,0% for model 9 to 20,0% for model 11. Using WS in dARX models (model 14), a sensitivity of 80,0% was reached. This value was almost 7% higher than the sensitivity of model 12 (73,3%) and approached the results found for the statistical models. The specificity of model 12 was equal to 52,0% and dropped to 20,0% for model 14. The values calculated for model 14 approached the results found for the statistical models.

The model with the highest sensitivity (82,9%) was model 6. This results confirms that model 6 was the best ALDM developed in this study.

6. Discussion

In this section the obtained results are discussed. An interpretation of the outcome of the experiment is given and suggestions towards future research are made. The ALDM that were developed are analysed: a possible explanation for the model structures and an interpretation of the parameters is given. The performance of the models is reviewed critically and compared with the results published in literature. The advantages and disadvantages of the ALDM developed in this study are discussed.

6.1 Interpretation effect lameness state on walking speed

The effect of lameness state on the manual walking speed of the cows was investigated using the MWS measurements. Due to the normalisation formula presented in Equation 1, an inversion took place. This inversion declares why the MWS measurements of lame cows are higher than the MWS measurements of non-lame cows as shown in Figure 4.

A significant difference in MWS was found from the ANOVA repeated measures analysis. The mean MWS of lame cows was equal to 0,91 s^2/m^2. This value was significantly higher than the mean MWS of non-lame cows (0,60 s^2/m^2). As explained in appendix 2 the normalisation procedure was mere statistical. The same formula was used for the normalisation of the original manual walking speed measurements of non-lame and lame cows. Therefore the significant difference found for the MWS measurements of non-lame and lame cows was also valid for the original manual walking speed measurements. The original manual walking speed of lame cows (1,05 m/s) was significantly lower than the value found for non-lame cows (1,29 m/s). The difference in walking speed was equal to 0,24 m/s. The values found by Chapinal et al. (2010) and Blackie et al. (2011b) were equal to 0,14 m/s and 0,55 m/s respectively. The difference in walking speed found in this study was intermediate to these values and approached the value of 0,30 m/s that was used as an estimate for the expected difference in the determination of the number of cows needed in the experiment. The assumption of the expected difference in walking speed between non-lame and lame cows was correct and so was the number of cows that was determined to ensure a statistical reliable experiment.

For the MWS measurements 94,3% of the total variance was related to differences between cows. The variance between the three measurements of one cow (the within cows variance) was limited to 5,7%. This last value can be seen as a measure for the reliability and the accuracy of the measurements. Since the three measurements for one cow were obtained within 90 minutes, the different walking speed measurements collected for one cow were not expected to differ drastically. Based on the calculated proportions of variance the MWS measurements and the results of the analysis using the MWS measurements were assumed reliable. Using manual measurements, a significant difference in walking speed between non-lame and lame cows was found. The first objective was proven successfully.

6.2 Interpretation collection automated walking speed measurements

As presented in Equation 2 the normalisation used for the original automated walking speed measurements did also involve an inversion. This inversion declares why higher AWS measurements were found for lame cows. The mean AWS was equal to 2,58 s/pixel for non-lame cows. The mean AWS of lame cows was higher (3,18 s/pixel). There is a trend of higher AWS measurements for lame cows, but the difference was not significant at the chosen significance level (as illustrated by the P-value of 0,066 in Table 11). As described for the MWS measurements in the previous section, the same conclusions were drawn for the original automated walking speed measurements and the normalised AWS measurements. The average original automated walking speed was equal to 0,38 pixels/s and 0,31 pixels/s for non-lame and lame cows respectively. Although a trend was visible, the difference that was found was not significant for the automated walking speed.

The within cows variance was about three times higher for the AWS than for the MWS measurements (18,8% compared to 5,7%). The between cows variance remained more than four times higher but the accuracy and the reliability of the AWS measurements decreased drastically compared to the MWS measurements.

The difference in walking speed that was found using manual measurements could not be identified using automated measurements. The second objective could not be confirmed based on the data used in this study. Two major influencing factors are suggested:

1) For the manual measurements the walking speed was calculated by dividing the length of the trajectory by the time needed to complete it. The resulting walking speed can therefore be seen as an average value over a distance of 16,43m. For the automated walking speed measurements the speed was calculated from the video recordings. Due to the limited field of view of the camera (3,74m at ground level), the automated walking speed measurements give an average value over a distance that is approximately 5 times shorter. Any disturbance or irregularity in the walking pattern of the cow that occurs within the field of view of the camera will affect the automated

walking speed measurements harder than the manual walking speed measurements since the buffer is about 5 times smaller.

2) For the analysis of the automated walking speed measurements only 27 cows were involved. Although 27 cows are sufficient for a reliable statistical analysis as illustrated in Table 1, the sample size is relatively small. Using small samples, every individual has an important effect on the outcome of the analysis. As can be seen in Figure 8 the AWS measurements for cow 8 are high if compared with the AWS measurements of the other non-lame cows. If this cow would not have been involved, there is a realistic possibility that the calculated P-value would have been smaller than the 5% significance level and that a significant difference was found.

On the day of the experiment, the difference in automated walking speed measurements was not significant for the non-lame and lame cows involved in the experiment. If walking speed measurements are collected on a regular base over a longer period, the change in and the difference between consecutive measurements can be used to investigate the effect of lameness state on walking speed.

For future research, the procedure that was used in this study to measure walking speed automatically could be altered in two ways. A first possibility is the use of additional video cameras. If a series of cameras (indicated by the red squares in Figure 13) is installed above the alley connecting the milking parlour with the housing pen, video recordings obtained from the different cameras could be combined to calculate one average value over a longer distance. The problems due to the limited field of view of one camera could be solved using this procedure. However, this solution requires more complex technical support because more cameras are involved. The computational power will also be higher if multiple cameras are involved.

Another possibility is to install a transmitter at the beginning and at the end of the alley. A possible position of the transmitters is suggested by the blue X sign in Figure 13. If a cow passes the first transmitter, a time measurement could be started. If the time needed to reach the second transmitter is measured and the distance between the two transmitters is known, the walking speed could be calculated as described in this study for the original manual walking speed measurements.

There is however a big difference in collecting walking speed measurements under experimental and under practical conditions. The longer the distance that is used to measure the walking speed, the bigger the chance that other cows will disturb the measurement. The number of cows in the alley at the same time has to be kept as low as possible. A gate, indicated by the vertical blue line in Figure 13, at the entry of the alley that opens as soon as the previous cow left the alley could be considered. It is important to avoid that waiting cows disturb the milking process. If the milking process is affected, the farmer will need more time to finish the milking. An 'after milking waiting area' between the exit of the milking parlour and the entry of the alley might be necessary to reduce the impact of the waiting cows.

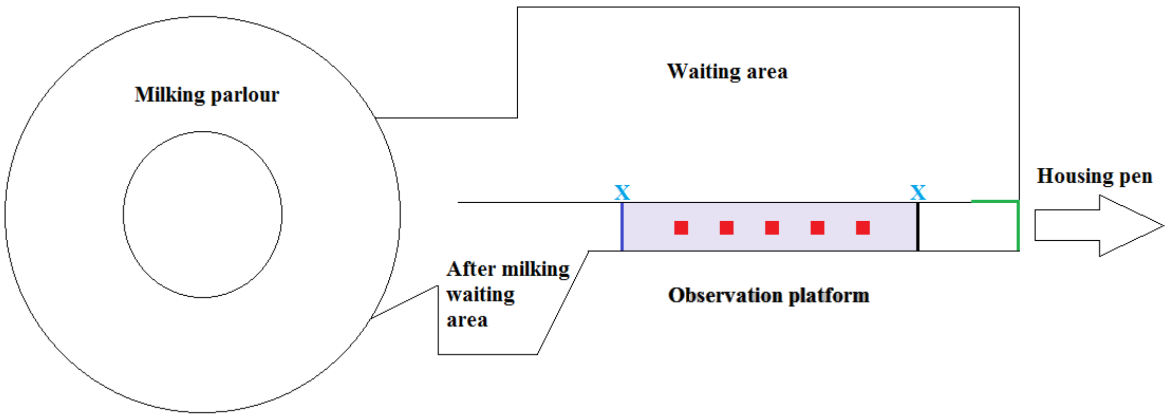

Figure 13 : Possible adaptations for future automated walking speed measurements. The observation platform from figure 1 is not presented here. The red squares indicate the positions of the different video cameras. The position of the transmitters is presented by the blue X signs and the vertical blue line shows the position of the gate at the entry of the alley. The green lines represent the selection gates.

6.3 Discussion development of ALDM

6.3.1 Remarks regarding data used for ALDM development

For the development of the ALDM, data collected between April 4^{th} and July 31^{st} were used. Two problems concerning the collected measurements were faced:

1) A substantial number of cows had missing data for at least one of the feature variables. Missing activity data can be declared by low batteries in the collar collecting the measurements. Drying off of cows led to missing values for back posture and walking speed. When cows were dried off, they were not milked and no video recordings were made. This problem has to be considered for the on-farm implementation of ALDM based on video recordings.
2) Cows in lactation passed through the alley twice a day. 14 recordings were made per cow every week. Due to useless recordings (e.g. missing identification, inferior quality of the video images,…) only a few correct recordings per cow were retained every week. Since reliable information was needed at every point in time, the correct

recordings of the last week were used to calculate one average value of back posture and walking speed. An average activity value over the last week was also calculated. Using this procedure, ACT, BPM, WS and BPM / WS were created and a reliable value for activity, back posture and walking speed was guaranteed at any point in time. The biggest disadvantage of this approach was that the dynamics of the parameters were also averaged over the last week. Adaptations as discussed in the previous section might increase the number of correct measurements. The period over which an average value has to be calculated can become smaller and a change in dynamics will be captured earlier.

6.3.2 Discussion of model structures

In the selected models ACT was almost not used. Only model 12 did use ACT as an input. The other selected models were based on BPM, WS and BPM / WS. For the statistical models the use of ACT as an extra variable did not increase the performance (model 2 was preferred over model 1 and model 6 was preferred over model 7). This was remarkable since the use of an additional explanatory variable was expected to increase the performance of the ALDM.

A possible explanation lies in the effects of all kinds of factors on the activity of the cows: different environmental, physiological and management related factors are known to influence the activity of the cows. One of the most striking factors is oestrus. According to Williamson et al. (2006) the activity of cows in heat can be up to four times higher than under normal conditions. In Figure 14 the daily activity of a lame cow between April 4^{th} and May 4^{th} 2014 is shown. Within this period peaks in activity were noticed at day 2 and day 23. The activity on those days was about two times higher than on other days. The difference of 21 days between the peaks suggest that the cow was in heat on those days and that heat influences the activity of lame cows. The effects of other, non-lameness related factors on the activity of the cows can declare why ACT was not very useful in the model types investigated in this study. Since heat is a regularly recurring factor until insemination is successful, it might be interesting to try to remove high fluctuations in the activity pattern due to heat in future research.

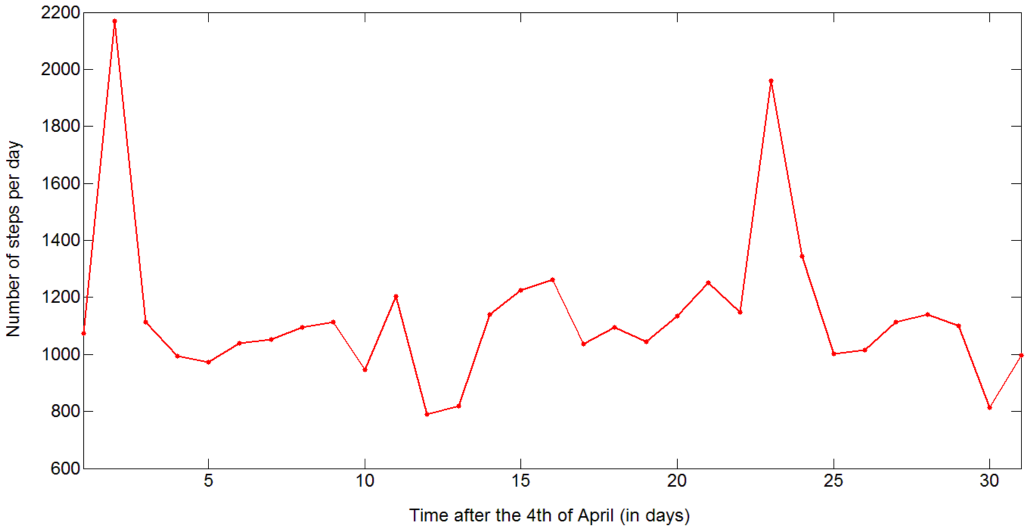

Figure 14 : The daily activity of a lame cow between April 4th and May 4th 2014. Two peaks, at day 2 and day 23 are noticed.

6.3.2.1 Statistical models

The range of the back posture and the automated walking speed of the cows was similar. In Equation 12 the BPM and WS coefficients were of similar order of magnitude as well. Based on the values of the coefficients, the weight assigned to the BPM and WS variables was similar. High lameness probabilities were found if BPM was high and WS was low.

In previous research (Viazzi et al., 2013, Van Hertem et al., 2014) the influence of lameness on back posture was described. Lame cows showed a higher back posture than their non-lame herdmates. This relationship was confirmed by the positive sign found for the BPM coefficient in Equation 12. The negative sign of the WS coefficient proved that WS and BPM were affected in opposite ways: the walking speed was lower for lame cows.

6.3.2.2 dAR models

On a commercial dairy farm the farmer is responsible for different tasks. Some are very general and take place on a regular base, other are really specific and only happen once or twice a year. The farmer develops his own routine to complete all the tasks and implements this routine in the farm management. The most regular tasks include the milking and feeding of the animals. On the farm visited in this study cows were fed a roughage mixture every day. Milking took place twice a day and the roughage mixture was pushed up twice every day as well. Bedding material in the cubicles was changed every few days. A more exceptional task for example was the professional claw trimming session that was organised on April 3rd and April 4th 2014.

Every task influences the behaviour of the cows. Most interventions only cause small alterations of the behaviour of the cows but others have a bigger impact (Cooper et al., 2008). The MW used for estimation had to include as many of the regular tasks as possible. By selecting an estimation MW of twelve days, the regular tasks were included in the dataset used for parameter estimation at any time. This also declares why the effect of an additional day was bigger for small values of the estimation MW.

As illustrated in Figure 10 the effect of the size of the prediction MW was not very clear for values between five and fifteen days. As long as the estimation MW included the general tasks a reliable estimation could be made for the entire period. Due to the nature of the animal and the characteristics of the problem (lameness evolves gradually) a drastic change was not expected within the period that is presented in Figure 10. If longer prediction periods would be considered, the mean squared errors are expected to increase. The optimal prediction MW was determined at seven days. This gives the farmer a few days to schedule the treatment in his routine but still guarantees an early treatment for the animal.

For the dAR models only the previous result was used in the calculation of the new parameter $a_{1.}$ This might seem strange but it is important to notice how the variables that were used in the development of the dAR models were constructed. As explained in Table 5 the ACT, BPM, WS, and BPM / WS values represent an average value over the last week. This means that each measurement contains information from the last week. Since the previous measurement was always involved in the new calculations and calculations were repeated every day, information from the last eight days was used in the new calculations. This once more confirms the importance to include the effects of the most regular tasks on a farm.

6.3.2.3 dARX models

The same optimal size of the MW used for estimation and prediction as described for the dAR models was found for the dARX models. The discussion regarding the size of the moving windows made in the previous section is also valid for the dARX models.

The optimal number of parameters involved in the dARX models was determined: 2 a parameters (a_1 and a_2) and one b parameter (b_0). Based on the same reasoning as discussed above the previous 2 measurements were involved in the new calculations. Information from the last nine days was used in each calculation confirming the needs to include the most regular tasks in the MW for estimation. The time delay was equal to zero. Given the construction of the ACT, BPM, WS and BPM / WS variables, this means that the input variable affects the output variable within the same week.

The b_0 parameter, corresponding with the input variable BPM in model 12 and model 14, was selected for the validation of the dARX models. This parameter corresponds with the input variable BPM for model 12 and model 14. The b_0 parameter was dominant over the other parameters and the lameness state affected the value of the parameter rather than the rate of change or the variance. The BPM variable was the most interesting variable as could be expected based on the ALDM that only use back posture information (Table 1 and Table 2). The interaction with another variable added some information, but the variable itself or the parameters corresponding with this variable were not used in the end.

6.3.3 Discussion of model performance

The model performance was expected to increase from statistical models over dAR models to dARX models. Compared to statistical models the performance of dAR models was expected to increase since the dynamics of the cow individual parameters were considered. The combination of different variables in dARX models was expected to lead to an additional increase in performance.

The statistical models provided a good starting point. Using the high specificity thresholds, specificities around 80% were obtained. The ALDM published in literature since 2002 reached specificities up to 96,4% (Schlageter-Tello et al., 2004). The sensitivities achieved using statistical models were limited to values around 20%. This is approximately 10% lower than the sensitivities that were found using manual locomotion scoring systems (Espejo et al., 2006, Fabian et al., 2014). The statistical models are group models. In these models each of the measurements was used independently and the individual dynamics were not considered. It was expected that this approach would not be sufficient to reach the targets specified in the third objective.

The use of dAR models, considering the dynamics of the parameters, did not to result in higher performances compared to the statistical models. Using the high specificity thresholds, specificities around 72% were reached. This was about 8% lower than the values found for the statistical models. The sensitivities were limited to values around 22%. The performances of the selected dAR models were lower than for the statistical models. An explanation can be found in the way the ACT, BPM, WS and BPM / WS variables were calculated. Due to the high number of useless recordings an average value had to be calculated for the variables.

This procedure provided a representative measure for ACT, BPM, WS and BPM / WS at any point but nullified the biggest advantage of the dAR models: the dynamics of the parameters got lost in the calculations of the average value.

Compared to the dAR models the performance of the dARX models was higher. The combination of different variables did have an effect. specificities of 76,4% and 79,5% were obtained and the sensitivities were equal to 24,1% and 22,3%. These values approach the results of the statistical models. Since the same setup and the same data preparation was used for the dAR and the dARX models, the dynamics were also lost in the dARX models. The increase in performance compared to the dAR models can be attributed to the interactions between different variables.

In the review presented by Schlageter et al. (2014) the specificities and sensitivities of seven ALDM developed between 2002 and 2013 were given. The specificities listed in that review varied between 72,8% and 96,4%. Sensitivities between 39,1% and 90% were achieved. The lowest specificities were found for models with high sensitivity values and vice versa. The results achieved by Van Hertem et al. (2014) were in the same range: the specificity was equal to 94,1% and a sensitivity of 54,9% was reached.

For the ALDM developed in this study, the best results were found for model 6. The specificity for this model was equal to 81,6% and a sensitivity of 18,3% was found. Similar results were found for model 14 but model 6 was easier and thus preferred. These values are worse than the results given in the review of Schlageter et al. (2014) or the study presented by Van Hertem et al. (2014) and are not good enough to implement the ALDM on a commercial dairy farm. Using the setup described in this study, the number of useful video recordings was insufficient to develop high-performing dAR and dARX models. Limitations of the modelling strategies, the setup and the processing of the variables are the most likely reasons why the performance of the developed models was lower than the results published in literature.

The best ALDM developed in this study was based on BPM and WS. Measurements for these variables were extracted from video recordings. Data were collected using a single commercial video camera for the entire herd. Cow individual sensors such as pedometers or accelerometers were not involved in these models. The use of a video camera is interesting for the farmer because the use of a general system based on one camera is cheaper than the use of cow individual sensors. Given the small profit margins in the dairy sector every contribution is important (St-Pierre, 2001).

For the cows a model based on video recordings is interesting because it is non-invasive (Van Hertem et al., 2013a). Although the effect of a pedometer or a neck collar on the behaviour of a cow is limited, it is always better if this is not necessary to reduce the impact on the behaviour of the cows.

A few disadvantages of ALDM based on video recordings have to be considered. As explained earlier it is often difficult to collect useful recordings on every day’s base. If the camera breaks or a problem occurs no information for any cow is available. If an individual sensor fails the loss of information stays limited to one cow. Another problem is that cows are separated from the herd if they are not in lactation. If cows do not pass underneath the video camera, no video recordings are made and no information is available.

7. Conclusion

A significant difference in walking speed was found using the manual measurements collected during the walking speed experiment. The manual walking speed of lame cows (1,05 m/s) was significantly lower than the manual walking speed of non-lame cows (1,29 m/s). Using the automated walking speed measurements extracted from the video recordings, the significant difference could not be confirmed. The automated walking speed of lame cows (0,31 pixels/s) was lower than the automated walking speed of non-lame cows (0,38 pixels/s) but the difference was no longer significant. Compared to the manual measurements, the within cows variance was about three times higher for the automated measurements. The accuracy of the automated measurements was limited.

Without WS a statistical model (model 2) with a specificity of 79,4% was developed. The sensitivity of this model was equal to 22,6%. Using dAR models the specificity decreased with about 8% compared to the statistical models. The performance of the best dARX model without WS (model 12) was similar to the performance of model 2: the specificity and the sensitivity were equal to 76,4% and 24,1% respectively. The results achieved for the specificity of the models were acceptable but the sensitivity values were not high enough.

Using WS as an extra feature variable the specificity of the statistical models increased with 2,2% to 81,6% for model 6. The sensitivity dropped with 4,3% to 18,3%. Similar trends were found for the dARX models were the use of WS led to a 3% increase in specificity and a 1,8% decrease in sensitivity compared to model 12. For model 14 the specificity was equal to 79,4% and a sensitivity of 22,3% was found. The results of the dAR models involving WS were insufficient reaching a maximal specificity of 72,1%. The use of WS did not lead to a 5% increase in specificity and sensitivity as targeted in the fourth objective. Model 6 was selected as the best ALDM developed in this study.

As specified in the third objective the final goal was to develop an ALDM with a specificity of 80%. The sensitivity had to be at least 30% according to the results obtained using manual systems. Based on these values the performance of model 6 was not high enough. The specificity of 81,6% was sufficient but the sensitivity of 18,3% was too limited to implement the ALDM on a commercial dairy farm.

8. References

Akaike, H. 1974. A New Look at the Statistical Model Identification. Automatic Control, IEEE Transactions 19(6):716-723.

Alban, L. 1995. Lameness in Danish dairy cows: frequency and possible risk factors. Preventive Veterinary Medicine 22(3):213-225.

Alsaaod, M., C. Romer, J. Kleinmanns, K. Hendriksen, S. Rose-Meierhofer, L. Plumer, and W. Buscher. 2012. Electronic detection of lameness in dairy cows through measuring pedometric activity and lying behavior. Applied Animal Behaviour Science 142(3-4):134-141.

Aoki, Y., M. Kamo, H. Kawamoto, J. Zhang, and A. Yamada. 2006. Changes in walking parameters of milking cows after hoof trimming. Animal Science Journal 77(1):103-109.

Archer, S., N. Bell, and J. Huxley. 2010a. Lameness in UK dairy cows: a review of the current status. In Practice 32(10):492-504.

Archer, S., M. Green, and J. Huxley. 2010b. Association between milk yield and serial locomotion score assessments in UK dairy cows. Journal of Dairy Science 93(9):4045-4053.

Barker, Z., J. Amory, J. Wright, R. Blowey, and L. Green. 2007. Management factors associated with impaired locomotion in dairy cows in England and Wales. Journal of Dairy Science 90(7):3270-3277.

Barker, Z. E., K. A. Leach, H. R. Whay, N. J. Bell, and D. C. J. Main. 2010. Assessment of lameness prevalence and associated risk factors in dairy herds in England and Wales. Journal of Dairy Science 93(3):932-941.

Bauman, D. E. and W. B. Currie. 1980. Partitioning of nutrients during pregnancy and lactation: a review of mechanisms involving homeostasis and homeorhesis. Journal of Dairy Science 63(9):1514-1529.

Bewley, J., R. Boyce, J. Hockin, L. Munksgaard, S. Eicher, M. Einstein, and M. Schutz. 2010. Influence of milk yield, stage of lactation, and body condition on dairy cattle lying behaviour measured using an automated activity monitoring sensor. Journal of Dairy Research 77(1):1-6.

Bicalho, R., S. Cheong, G. Cramer, and C. Guard. 2007. Association between a visual and an automated locomotion score in lactating holstein cows. Journal of Dairy Science 90(7):3294-3300.

Bicalho, R., L. Warnick, and C. Guard. 2008. Strategies to analyze milk losses caused by diseases with potential incidence throughout the lactation: A lameness example. Journal of Dairy Science 91(7):2653-2661.

Bikker, J., H. van Lear, R. Rump, J. Doorenbos, K. van Meurs, G. Griffioen, and J. Dijkstra. 2014. Technical note: Evaluation of an ear-attached movement sensor to record cow feeding behavior and activity. Journal of Dairy Science 97(5):2974-2979.

Blackie, N., J. Amory, E. Bleach, and J. Scaife. 2011a. The effect of lameness on lying behaviour of zero grazed Holstein dairy cattle. Applied Animal Behaviour Science 134(3-4):85-91.

Blackie, N., E. Bleach, J. Amory, and J. Scaife. 2011b. Impact of lameness on gait characteristics and lying behaviour of zero grazed dairy cattle in early lactation. Applied Animal Behaviour Science 129(2-4):67-73.

Booth, C., L. Warnick, Y. Grohn, D. Maizon, C. Guard, and D. Janssen. 2004. Effect of lameness on culling in dairy cows. Journal of Dairy Science 87(12):4115-4122.

Borderas, T., A. Fournier, J. Rushen, and A. De Passille. 2008. Effect of lameness on dairy cows' visits to automatic milking systems. Canadian Journal of Animal Science 88(1):1-8.

Borderas, T., B. Pawluczuk, A. de Passille, and J. Rushen. 2004. Claw hardness of dairy cows: Relationship to water content and claw lesions. Journal of Dairy Science 87(7):2085-2093.

Brehme, U., U. Stollberg, R. Holz, and T. Schleusener. 2008. ALT pedometer - New sensor-aided measurement system for improvement in oestrus detection. Computers and Electronics in Agriculture 62(1):73-80.

Bruijnis, M., H. Hogeveen, and E. Stassen. 2010. Assessing economic consequences of foot disorders in dairy cattle using a dynamic stochastic simulation model. Journal of Dairy Science 93(6):2419-2432.

Chapinal, N., A. de Passille, J. Rushen, and S. Wagner. 2010. Automated methods for detecting lameness and measuring analgesia in dairy cattle. Journal of Dairy Science 93(5):2007-2013.

Clarkson, M., D. Downham, W. Faull, J. Hughes, F. Manson, J. Merritt, R. Murray, W. Russell, J. Sutherst, and W. Ward. 1996. Incidence and prevalence of lameness in dairy cattle. Veterinary Record 138(23):563-567.

Cook, N. and K. Nordlund. 2009. The influence of the environment on dairy cow behavior, claw health and herd lameness dynamics. Veterinary Journal 179(3):360-369.

Cooper, M., D. Arney, and C. Phillips. 2008. The effect of temporary deprivation of lying and feeding on the behaviour and production of lactating dairy cows. Animal 2(2):275-283.

Cramer, G., K. Lissemore, C. Guard, K. Leslie, and D. Kelton. 2008. Herd- and cow-level prevalence of foot lesions in Ontario dairy cattle. Journal of Dairy Science 91(10):3888-3895.

de Mol, R., G. Andre, E. Bleumer, J. van der Werf, Y. de Haas, and C. van Reenen. 2013. Applicability of day-to-day variation in behavior for the automated detection of lameness in dairy cows. Journal of Dairy Science 96(6):3703-3712.

Diskin, M. and J. Sreenan. 2000. Expression and detection of oestrus in cattle. Reproduction Nutrition Development 40(5):481-491.

Dohoo, I. R. and S. W. Martin. 1984. Disease, production and culling in Holstein-Friesian cows. Preventive Veterinary Medicine 2(6):755-770.

Enting, H., D. Kooij, A. Dijkhuizen, R. Huirne, and E. NoordhuizenStassen. 1997. Economic losses due to clinical lameness in dairy cattle. Livestock Production Science 49(3):259-267.

Espejo, L., M. Endres, and J. Salfer. 2006. Prevalence of lameness in high-producing Holstein cows housed in freestall barns in Minnesota. Journal of Dairy Science 89(8):3052-3058.

Fabian, J., R. Laven, and H. Whay. 2014. The prevalence of lameness on New Zealand dairy farms: A comparison of farmer estimate and locomotion scoring. Veterinary Journal 201(1):31-38.

Fenwick, M. A., S. Llewellyn, R. Fitzpatrick, D. A. Kenny, J. J. Murphy, J. Patton, and D. C. Wathes. 2008. Negative energy balance in dairy cows is associated with specific changes in IGF-binding protein expression in the oviduct. Society for Reproduction and Fertility 135(1):63-75.

Flower, F., A. de Passille, D. Weary, D. Sanderson, and J. Rushen. 2007. Softer, higher-friction flooring improves gait of cows with and without sole ulcers. Journal of Dairy Science 90(3):1235-1242.

Flower, F., D. Sanderson, and D. Weary. 2005. Hoof pathologies influence kinematic measures of dairy cow gait. Journal of Dairy Science 88(9):3166-3173.

Flower, F., D. Sanderson, and D. Weary. 2006. Effects of milking on dairy cow gait. Journal of Dairy Science 89(6):2084-2089.

Flower, F. and D. Weary. 2006. Effect of hoof pathologies on subjective assessments of dairy cow gait. Journal of Dairy Science 89(1):139-146.

Flower, F. and D. Weary. 2009. Gait assessment in dairy cattle. Animal 3(1):87-95.

Fregonesi, J., C. Tucker, and D. Weary. 2007. Overstocking reduces lying time in dairy cows. Journal of Dairy Science 90(7):3349-3354.

Garbarino, E., J. Hernandez, J. Shearer, C. Risco, and W. Thatcher. 2004. Effect of lameness on ovarian activity in postpartum Holstein cows. Journal of Dairy Science 87(12):4123-4131.

Gregory, N., L. Craggs, N. Hobson, and C. Krogh. 2006. Softening of cattle hoof soles and swelling of heel horn by environmental agents. Food and Chemical Toxicology 44(8):1223-1227.

Hall, S. J. 1995. Basic Biomechanics. 2nd ed. William C Brown Publishers, Dubuque, Iowa , US.

Haskell, M. J., L. J. Rennie, V. A. Bowell, M. J. Bell, and A. B. Lawrence. 2006. Housing System, Milk Production, and Zero-Grazing Effects on Lameness and Leg Injury in Dairy Cows. Journal of Dairy Science 89(11):4259-4266.

Herlin, A. H. and S. Drevemo. 1994. Biomechanics of dairy cows; effects of different management systems. Sverige.

Hollenbeck, A. R. 1978. Problems of reliability in observational research. Pages 79-98 in Observing behaviour, vol. II. Data collection and analysis methods. Vol. 2. G. Sackett, ed. University Park Press, Baltimore.

Huybrechts, T., K. Mertens, J. De Baerdemaeker, B. De Ketelaere, and W. Saeys. 2014. Early warnings from automatic milk yield monitoring with online synergistic control. Journal of Dairy Science 97(6):3371-3381.

Ito, K., M. von Keyserlingk, S. LeBlanc, and D. Weary. 2010. Lying behavior as an indicator of lameness in dairy cows. Journal of Dairy Science 93(8):3553-3560.

Juarez, S., P. Robinson, E. DePeters, and E. Price. 2003. Impact of lameness on behavior and productivity of lactating Holstein cows. Applied Animal Behaviour Science 83(1):1-14.

Junge, W. 1997. Effects on the health of claws of dairy cows. Züchtungskunde 69(2):122-129.

Kamphuis, C., E. Frank, J. Burke, G. Verkerk, and J. Jago. 2013. Applying additive logistic regression to data derived from sensors monitoring behavioral and physiological characteristics of dairy cows to detect lameness. Journal of Dairy Science 96(11):7043-7053.

Keady, T., C. Mayne, D. Fitzpatrick, and M. McCoy. 2001. Effect of concentrate feed level in late gestation on subsequent milk yield, milk composition, and fertility of dairy cows. Journal of Dairy Science 84(6):1468-1479.

Klaas, I., T. Rousing, C. Fossing, J. Hindhede, and J. Sorensen. 2003. Is lameness a welfare problem in dairy farms with automatic milking systems? Animal Welfare 12(4):599-603.

Kutner, M. H., C. J. Nachtsheim, J. Neter, and W. Li. 2005. Applied Linear Statistical Models. 5th ed. McGraw-Hill/Irwin.

Leach, K., D. Tisdall, N. Bell, D. Main, and L. Green. 2012. The effects of early treatment for hindlimb lameness in dairy cows on four commercial UK farms. Veterinary Journal 193(3):626-632.

Mackinson, R., J. Bewley, G. Heersche, and B. Silvia. 2013. Automated Heat Detection May Change The Way You Breed Cows. Vol. 2014. University of Kentucky, Kentucky.

Maertens, W., J. Vangeyte, J. Baert, A. Jantuan, K. Mertens, S. De Campeneere, A. Pluk, G. Opsomer, S. Van Weyenberg, and A. Van Nuffel. 2011. Development of a real time

cow gait tracking and analysing tool to assess lameness using a pressure sensitive walkway: The GAITWISE system. Biosystems Engineering 110(1):29-39.
Main, D., J. Clegg, A. Spatz, and L. Green. 2000. Repeatability of a lameness scoring system for finishing pigs. Veterinary Record 147(20):574-576.
Manske, T., J. Hultgren, and C. Bergsten. 2002. Prevalence and interrelationships of hoof lesions and lameness in Swedish dairy cows. Preventive Veterinary Medicine 54(3):247-263.
Manson, F. and J. Leaver. 1988. The influence of concentrate amount on locomotion and clinical lameness in dairy cattle. Animal Production 47:185-190.
March, S., J. Brinkmann, and C. Winkler. 2007. Effect of training on the inter-observer reliability of lameness scoring in dairy cattle. Animal Welfare 16(2):131-133.
Marquer, P. 2014. Milk and dairy production statistics. Vol. 2014. Eurostat, ed. European Commission.
Martin, S., S. Aziz, W. Sandals, and R. Curtis. 1982. The association between clinical disease, production and culling of Holstein-Friesian cows. Canadian Journal of Animal Science 62(2):633-640.
O'Callaghan, K., P. Cripps, D. Downham, and R. Murray. 2003. Subjective and objective assessment of pain and discomfort due to lameness in dairy cattle. Animal Welfare 12(4):605-610.
O'Driscoll, K., L. Boyle, and A. Hanlon. 2008. A brief note on the validation of a system for recording lying behaviour in dairy cows. Applied Animal Behaviour Science 111(1-2):195-200.
Opsomer, G., M. Coryn, H. Deluyker, and A. de Kruif. 1998. An analysis of ovarian dysfunction in high yielding dairy cows after calving based on progesterone profiles. Reproduction in Domestic Animals 33(3-4):193-204.
Pastell, M. and M. Kujala. 2007. A probabilistic neural network model for lameness detection. Journal of Dairy Science 90(5):2283-2292.
Pastell, M., H. Takko, H. Grohn, M. Hautala, V. Poikalainen, J. Praks, I. Veermae, M. Kujala, and J. Ahokas. 2006. Assessing cows' welfare: weighing the cow in a milking robot. Biosystems Engineering 93(1):81-87.
Pastell, M., J. Tiusanen, M. Hakojarvi, and L. Hanninen. 2009. A wireless accelerometer system with wavelet analysis for assessing lameness in cattle. Biosystems Engineering 104(4):545-551.
Peham, C., T. Licka, D. Girtler, and M. Scheidl. 2001. Hindlimb lameness: clinical judgement versus computerised symmetry measurement. Veterinary Record 148(24):750-752.
Platteau, J., D. Van Gijseghem, T. Van Bogaert, and E. Maertens. 2012. Landbouwrapport 2012. D. L. e. Visserij, ed. Jules Van Liefferinge, Brussel.
Pluk, A., C. Bahr, A. Poursaberi, W. Maertens, A. van Nuffel, and D. Berckmans. 2012. Automatic measurement of touch and release angles of the fetlock joint for lameness detection in dairy cattle using vision techniques. Journal of Dairy Science 95(4):1738-1748.
Potterton, S., N. Bell, H. Whay, E. Berry, O. Atkinson, R. Dean, D. Main, and J. Huxley. 2012. A descriptive review of the peer and non-peer reviewed literature on the treatment and prevention of foot lameness in cattle published between 2000 and 2011. Veterinary Journal 193(3):612-616.
Poursaberi, A., C. Bahr, A. Pluk, A. Van Nuffel, and D. Berckmans. 2010. Real-time automatic lameness detection based on back posture extraction in dairy cattle: Shape analysis of cow with image processing techniques. Computers and Electronics in Agriculture 74(1):110-119.

Rajala-Schultz, P., Y. Grohn, and C. McCulloch. 1999. Effects of milk fever, ketosis, and lameness on milk yield in dairy cows. Journal of Dairy Science 82(2):288-294.

Rajkondawar, P., A. Lefcourt, N. Neerchal, R. Dyer, M. Varner, B. Erez, and U. Tasch. 2002a. The development of an objective lameness scoring system for dairy herds: Pilot study. Transactions of the Asae 45(4):1123-1125.

Rajkondawar, P., M. Liu, R. Dyer, N. Neerchal, U. Tasch, A. Lefcourt, B. Erez, and M. Varner. 2006. Comparison of models to identify lame cows based on gait and lesion scores, and limb movement variables. Journal of Dairy Science 89(11):4267-4275.

Rajkondawar, P., U. Tasch, A. Lefcourt, B. Erez, R. Dyer, and M. Varner. 2002b. A system for identifying lameness in dairy cattle. Applied Engineering in Agriculture 18(1):87-96.

Renn, N., J. Onyango, and W. McCormick. 2014. Digital Infrared Thermal Imaging and manual lameness scoring as a means for lameness detection in cattle. Veterinary Clinical Science 2(2):16-23.

Rouha-Mulleder, C., C. Iben, E. Wagner, G. Laaha, J. Troxler, and S. Waiblinger. 2009. Relative importance of factors influencing the prevalence of lameness in Austrian cubicle loose-housed dairy cows. Preventive Veterinary Medicine 92(1-2):123-133.

Rutten, C. J., A. G. J. Velthuis, W. Steeneveld, and H. Hogeveen. 2012. Invited review: Sensors to support health management on dairy farms. Journal of Dairy Science 96(4):1928-1952.

Schlageter-Tello, A., E. Bokkers, P. Koerkamp, T. Van Hertem, S. Viazzi, C. Romanini, I. Halachmi, C. Bahr, D. Berckmans, and K. Lokhorst. 2014. Manual and automatic locomotion scoring systems in dairy cows: A review. Preventive Veterinary Medicine 116(1-2):12-25.

Scott, G. 1988. Studies of the gait of Friesian heifer cattle. Veterinary Record 123(10):245-248.

Scott, G. 1989. Changes in limb loading with lameness for a number of Friesian cattle. British Veterinary Journal 145(1):28-38.

Somers, J. G. C. J., K. Frankena, E. N. Noordhuizen-Stassen, and J. H. M. Metz. 2005. Risk factors for interdigital dermatitis and heel erosion in dairy cows kept in cucible houses in The Netherlands. Preventive Veterinary Medicine 71(1-2):23-34.

Song, X., T. Leroy, E. Vranken, W. Maertens, B. Sonck, and D. Berckmans. 2008. Automatic detection of lameness in dairy cattle - Vision-based trackway analysis in cow's locomotion. Computers and Electronics in Agriculture 64(1):39-44.

Sprecher, D., D. Hostetler, and J. Kaneene. 1997. A lameness scoring system that uses posture and gait to predict dairy cattle reproductive performance. Theriogenology 47(6):1179-1187.

St-Pierre, N. R. 2001. Invited review: integrating quantitative findings from multiple studies using mixed model methodology. Journal of Dairy Science 84(4):741-755.

Tadich, N., E. Flor, and L. Green. 2010. Associations between hoof lesions and locomotion score in 1098 unsound dairy cows. Veterinary Journal 184(1):60-65.

van der Tol, P., J. Metz, E. Noordhuizen-Stassen, W. Back, C. Braam, and W. Weijs. 2002. The pressure distribution under the bovine claw during square standing on a flat substrate. Journal of Dairy Science 85(6):1476-1481.

van Eerdenburg, F. J. C. M., H. de Boer, and F. Gómez. 2003. Relationship among subclinical lameness and oestrus expression. Pages 211-216 in Proceedings of the 11th International Congress of the International Society of Animal Hygiene. Mexico City.

Van Hertem, T., V. Alchanatis, A. Antler, E. Maltz, I. Halachmi, A. Schlageter-Tello, C. Lokhorst, S. Viazzi, C. Romanini, A. Pluk, C. Bahr, and D. Berckmans. 2013a.

Comparison of segmentation algorithms for cow contour extraction from natural barn background in side view images. Computers and Electronics in Agriculture 91:65-74.

Van Hertem, T., E. Maltz, A. Antler, C. Romanini, S. Viazzi, C. Bahr, A. Schlageter-Tello, C. Lokhorst, D. Berckmans, and I. Halachmi. 2013b. Lameness detection based on multivariate continuous sensing of milk yield, rumination, and neck activity. Journal of Dairy Science 96(7):4286-4298.

Van Hertem, T., S. Viazzi, M. Steensels, E. Maltz, A. Antler, V. Alchanatis, A. Schlageter-Tello, K. Lokhorst, E. Romanini, C. Bahr, D. Berckmans, and I. Halachmi. 2014. Automatic lameness detection based on consecutive 3D-video recordings. Biosystems Engineering 119:108-116.

Van Nuffel, A., J. Vangeyte, K. Mertens, L. Pluym, S. De Campeneere, W. Saeys, G. Opsomer, and S. Van Weyenberg. 2013. Exploration of measurement variation of gait variables for early lameness detection in cattle using the GAITWISE. Livestock Science 156(1-3):88-95.

van Winsen, F., Y. de Mey, E. Wauters, L. Lauwers, S. Van Passel, and M. Vancauteren. 2011. ILVO-mededeling nr. 84: Een volatiele melkprijs: het effect op het risicoprofiel van melkveebedrijven. I. v. L.-e. Visserijonderzoek, ed, Merelbeke.

Versonnen, A. 2012. Kerncijfers landbouw, de landbouw in België in cijfers. K. M. O. FOD Economie, Middenstand en Energie, ed, Brussel.

Viazzi, S., C. Bahr, A. Schlageter-Tello, T. Van Hertem, C. Romanini, A. Pluk, I. Halachmi, C. Lokhorst, and D. Berckmans. 2013. Analysis of individual classification of lameness using automatic measurement of back posture in dairy cattle. Journal of Dairy Science 96(1):257-266.

Viazzi, S., C. Bahr, T. Van Hertem, A. Schlageter-Tello, C. Romanini, I. Halachmi, C. Lokhorst, and D. Berckmans. 2014. Comparison of a three-dimensional and two-dimensional camera system for automated measurement of back posture in dairy cows. Computers and Electronics in Agriculture 100:139-147.

Warnick, L., D. Janssen, C. Guard, and Y. Grohn. 2001. The effect of lameness on milk production in dairy cows. Journal of Dairy Science 84(9):1988-1997.

Weary, D., L. Niel, F. Flower, and D. Fraser. 2006. Identifying and preventing pain in animals. Applied Animal Behaviour Science 100(1-2):64-76.

Welsh, E., G. Gettinby, and A. Nolan. 1993. Comparison of a visual analog scale and a numerical rating scale for assessment of lameness, using sheep as a model. American Journal of Veterinary Research 54(6):976-983.

Whay, H. 2002. Locomotion scoring and lameness detection in dairy cattle. In Practice 24(8):444-449.

Whay, H., D. Main, L. Green, and A. Webster. 2003. Assessment of the welfare of dairy cattle using animal-based measurements: direct observations and investigation of farm records. Veterinary Record 153(7):197-202.

Whay, H., A. Waterman, A. Webster, and J. O'Brien. 1998. The influence of lesion type on the duration of hyperalgesia associated with hindlimb lameness in dairy cattle. Veterinary Journal 156(1):23-29.

Whay, H. R., A. E. Waterman, and A. J. F. Webster. 1997. Associations between locomotion, claw lesions and nociceptive treshold in dairy heifers during the peri-partum period. Veterinary Journal 154(2):155-161.

Williamson, J. A., D. Bailey, and K. Butler. 2006. Electronic Heat Detection. in Proc. South Island Dairy Event. Massey University, New Zealand.

Winckler, C. and S. Willen. 2001. The reliability and repeatability of a lameness scoring system for use as an indicator of welfare in dairy cattle. Acta Agriculturae Scandinavica A: Animal Science 51(30):103-107.

Appendices

Appendix 1: transformation units

The automated walking speed measurements were expressed in pixels per second. The physically relevant unit for walking speed is meters per second. The relationship between the manual and the automated walking speed measurements collected during the experiment was investigated. A good relationship would have been interesting to transform the units of the automated walking speed measurements. As presented in Figure 15 the accuracy of the linear fit was limited. An R^2 value of 62,94% was found. The linear model is presented in Equation 15.

$$Manual\ walking\ speed\ = 2{,}0887 * Automated\ walking\ speed + 0{,}4212 \qquad \text{Equation 15}$$

The accuracy of the linear model was not high enough to transform the units of the automated walking speed measurements. The automated measurements were analysed using the pixels per second unit.

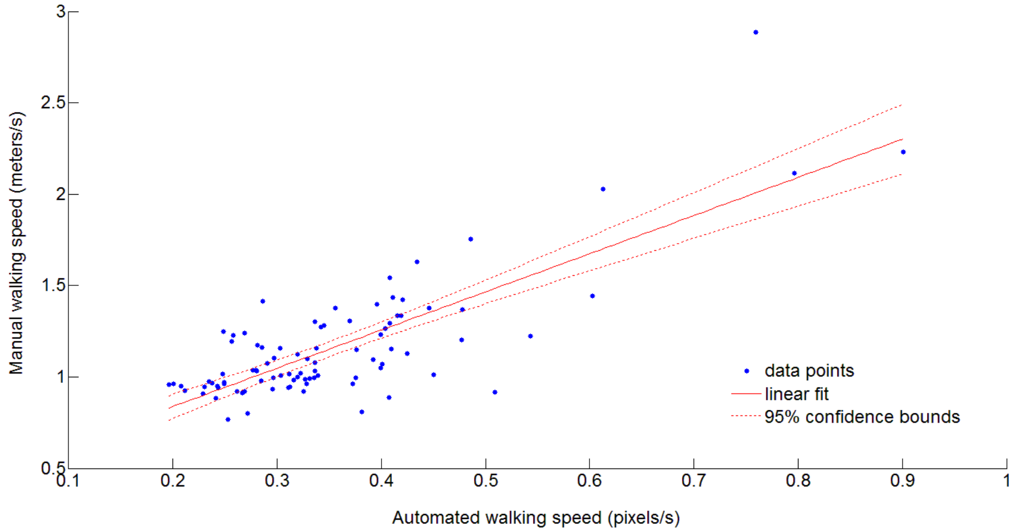

Figure 15 : Relationship between the manual (y-axis) and the automated (x-axis) walking speed measurements. The blue points represent the actual measurements. The red line indicates the linear fit and the red dotted lines give the 95% confidence bounds for the linear fit.

Appendix 2: Kruskal-Wallis analysis

The manual and the automated walking speed measurements collected during the experiment were normalised. This was needed because the ANOVA repeated measures analysis required the data to originate from a normally distributed population. The transformations presented in Equation 1 and Equation 2 were mere statistical and did not influence the outcome of the analysis.

This was illustrated using a Kruskal-Wallis test (Kutner et al., 2005). This test is the non-parametric alternative for the single factor ANOVA analysis. Since it is a non-parametric method, the Kruskal-Wallis test does not require the data to originate from a normally distributed population. The original manual and automated walking speed measurements were analysed and the results were compared with the outcome of the ANOVA repeated measures analysis. Eventually the ANOVA repeated measures analysis was preferred since the robustness of this test is higher than the robustness of the Kruskal-Wallis test.

The average manual walking speed, over the three repetitions, was calculated for each cow. The same procedure was used to calculate the average automated walking speed for cows with three correct measurements. The average values were investigated using the Kruskal-Wallis test and the results are presented in Table 16.

Table 16 : Results of the Kruskal-Wallis test for the walking speed measurements. The results are given for the manual and the automated walking speed measurements. Significant P-values are indicated with an asterisk.

	df	X^2 value	P-value
Manual walking speed measurements	1	4,978	0,026*
Automated walking speed measurements	1	3,768	0,052

Comparing the P-values given in the last column of Table 16 with the P-values calculated in Table 9 and Table 11, the same results were found. A significant difference in walking speed was found using the manual walking speed measurements. Based on the automated walking speed measurements the difference was no longer significant. This proves that the normalisation procedure did not affect the outcome of the analysis.

Appendix 3 : Overview developed models

The 14 ALDM that were developed in this study are presented in Table 17. This table gives a general overview of the ALDM. For each model the model type and the final feature variables that were involved in the development of the model can be found. For the dARX models a distinction between the input and the output of the models iss made.

Table 17 : Overview of the 14 ALDM developed in this study. The model type and the final feature variables used in each model are given. For the dARX models a distinction between the input and the output is made.

	Model type				**Final feature variables**			
Model name	Statistical model	dAR model	dARX model		ACT	BPM	WS	BPM / WS
Model 1	x				x			
Model 2	x					x		
Model 3	x				x	x		
Model 4	x						x	
Model 5	x				x		x	
Model 6	x					x	x	
Model 7	x				x	x	x	
Model 8		x			x			
Model 9		x				x		
Model 10		x					x	
Model 11		x						x
Model 12			x	input		x		
				output	x			
Model 13			x	input				x
				output	x			
Model 14			x	input		x		
				output			x	