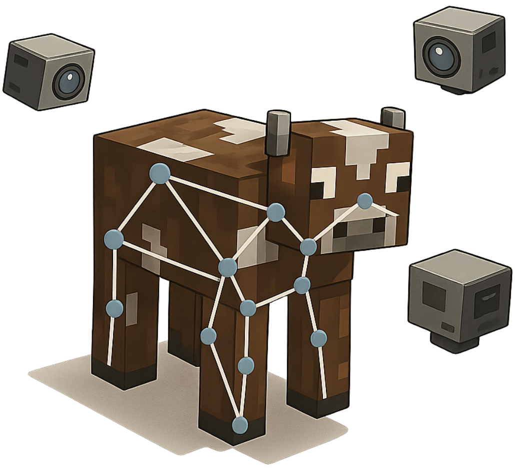
Acta Universitatis Agriculturae Sueciae

Doctoral Thesis No. 2026:6
Faculty of Veterinary Medicine and Animal Science

Dairy cow welfare & 3D computer vision

Evaluating cubicle comfort when getting up and lying down using pose estimation in 3D

Adrien Kroese



Dairy cow welfare & 3D computer vision

Evaluating cubicle comfort when getting up and lying down using pose estimation in 3D

Adrien Kroese
Faculty of Veterinary Medicine and Animal Science
Department of Clinical Sciences
Uppsala

SWEDISH UNIVERSITY
OF AGRICULTURAL
SCIENCES

DOCTORAL THESIS
Uppsala 2026

Acta Universitatis Agriculturae Sueciae
2026.6

Cover: "*A Minecraft cow being scanned by three cameras placed around it generating pose estimation"* Created with Dall-e3.

ISSN 1652-6880
ISBN (print version) 978-91-8124-203-4
ISBN (electronic version) 978-91-8124-223-2
https://doi.org/10.54612/a.54p7dcvs0a

Swedish University of Agricultural Sciences, Department of Clinical Sciences, Uppsala, Sweden
 Print: SLU Grafisk service, Uppsala 2026

Dairy cow welfare & 3D computer vision

Evaluating cubicle comfort when getting up and lying down using pose estimation in 3D

Abstract

This thesis sought to apply recent advances in precision livestock farming to evaluating a specific welfare parameter: cows' comfort when getting up and lying down in cubicles. Cows need sufficient space to get up and lie down but rigid metal bars interfere with their innate motion patterns. A multi-camera system recorded over 12 cubicles during two data-collection phases. Triangulating 24 anatomical landmarks detected on each view using computer vision produced 3D pose estimation throughout posture transitions. From these, the timing, spatial use, and weight distribution could be measured or modelled. When compared with human annotations, the system showed high agreement in identifying rising and lying down events and their phases. While stage detection was not fully repeatable, the results show that 3D keypoint motion reliably reflects observable kinematic patterns. The method developed was then applied to evaluate whether replacing metal head and neck rails with flexible straps improved cows' movement opportunities. An experiment was run in which the head and neck rails of cubicles were replaced with flexible straps. In the flexible setup, cows showed greater head vertical displacement and straighter lunge during rising, indicating greater movement opportunities. Effect sizes were small. Lying-down movements showed no consistent difference between flexible cubicles and both baselines. The duration of lying down movements increased upon returning to baseline, suggesting that duration alone doesn't fully capture comfort. There was a consistent difference in a novel indicator introduced in this work: the shift of the cows' centre of mass. The thesis concludes on the following: (i) Posture transition behaviours consist of multiple, independent dimensions and single-indicator assessments may not be a sound summarisation. (ii) Pose estimation in 3D represents a valuable technology to simultaneously monitor several indicators and uncover different strategies used to cope with restrictive environments. Finally (iii) that novel indicators such as modelled weight displacement are adapted to pose data and get a step closer to biomechanical drivers behind the specific motions and behaviours.

Keywords: precision livestock farming, animal welfare, free stall, rising behaviour, lying down behaviour, computer vision, dairy cattle, pose estimation

Välfärdsbedömning hos mjölkkor och 3D datorseende

Sammanfattning

Forskningen syftade till att tillämpa precisionsdjurhållning för att utvärdera en specifik välfärdsparameter: kors komfort när de reser sig och ligger sig ned i liggbåset. Kor behöver tillräckligt med utrymme och metallstänger kan störa deras naturliga rörelsemönster. Kameror användes för att filma 12 bås. Triangulering av 24 anatomiska landmärken som detekterades i varje vy med hjälp av datorseende gav en 3D-uppskattning av kroppsställningen under hela rörelseövergångarna. Utifrån dessa kunde tidpunkten, utrymmesanvändningen och viktfördelningen mätas eller modelleras. Jämfört med mänskliga annoteringar visade systemet hög överensstämmelse gällande att identifiera händelser när korna reste sig och lade sig, samt läggnings- och resningsbeteendets olika faser. Även om det inte gick att detektera stegen fullt ut, visar resultaten att 3D-nyckelpunktsrörelser på ett tillförlitligt sätt återspeglar observerbara mönster. Den utvecklade metoden användes sedan för att utvärdera om korna fick bättre rörelsefrihet när metallstängerna för huvud och hals ersattes med flexibla remmar. 3D-positioner samlades in under tre separata perioder om två veckor vardera, i baskonfigurationen, med flexibla huvud- och nackbommar och återigen i baskonfigurationen. I den flexibla konfigurationen visade korna större vertikal förskjutning av huvudet och rakare vinklar vid huvudutfallet under uppresningen, vilket indikerar större rörelsefrihet. Effektstorlekarna var dock små. Läggningsrörelserna visade ingen konsekvent skillnad mellan flexibla bås och båda baskonfigurationerna. Läggningsrörelsernas varaktighet ökade när man återgick till baslinjen, vilket tyder på att varaktigheten i sig inte helt återspeglar komforten. Avhandlingen drar följande slutsatser: (i) Beteenden vid ställningsförändringar består av flera oberoende dimensioner och bedömningar baserade på en enda indikator är kanske inte en tillförlitlig sammanfattning. (ii) Poseringsuppskattning i 3D är en värdefull teknik för att samtidigt övervaka flera indikatorer och upptäcka olika strategier som används för att hantera begränsande miljöer. Slutligen (iii) att nya indikatorer såsom modellerad viktförskjutning är anpassade till poseringsdata och kommer ett steg närmare de biomekaniska drivkrafterna bakom specifika rörelser och beteenden.

Evaluation du bien-être en bovins laitiers & vision par ordinateur

Résumé

Les mouvements de levers et couchers des bovins requièrent un espace suffisant, mais les barres métalliques peuvent entraver la cinématique naturelle des animaux. Un système de caméras a enregistré douze logettes en deux phases. Vingt-quatre repères anatomiques ont été détectés sur chaque vue par vision par ordinateur, puis triangulés afin d'estimer la posture en 3D tout au long des levers et couchers. Cette approche a permis de quantifier ou de modéliser le déroulement temporel des mouvements, l'utilisation de l'espace et la répartition du poids. Comparée aux annotations humaines, la méthode a montré une forte concordance pour l'identification des événements de lever et de coucher ainsi que de leurs phases. Bien que la détection fine des étapes ne soit pas parfaitement reproductible, les trajectoires 3D des points clés reflètent de manière fiable les schémas cinématiques observables. La méthode a ensuite été appliquée pour évaluer l'impact du remplacement des barres métalliques de tête et de cou par des sangles flexibles. Les poses en 3D ont été collectées sur trois périodes successives: configuration standard, configuration flexible, puis retour à la configuration standard. Avec des sangles flexibles, les vaches ont motré une plus grande amplitude de movements verticaux de la tête et des angles de fente plus droits lors du lever, suggérant une liberté de mouvement accrue, bien que les effets observés soient de faible amplitude. Les mouvements de coucher n'ont montré aucune différence syst´ematique entre les configurations. En revanche, la durée du couchage a augmenté lors du retour à la configuration standard, indiquant que la durée seule ne constitue pas un indicateur exhaustif du confort. Une différence constante a toutefois été observée pour un nouvel indicateur introduit: le déplacement du centre de gravité. La thèse conclut que : (i) les transitions posturales sont multidimensionnelles et ne peuvent être résumées de manière fiable par un indicateur unique ; (ii) l'estimation de la pose en 3D constitue un outil pertinent pour le suivi simultané de plusieurs indicateurs et pour l'identification de stratégies d'adaptation à des environnements contraignants ; et (iii) des indicateurs dérivés, tels que le déplacement du poids modélisé, sont adaptés aux données de posture et permettent d'accéder plus directement aux mécanismes biomécaniques sous-jacents aux mouvements et aux comportements observés.

Dedication

"*The dairy cow is entitled to be treated kindly*" (Cook & Nordlund, 2009)

There is an inherent paradox to welfare, isn't there? We seek to ease the burden created by the conditions that we ourselves impose on the animals. The progress is slow; each insight recorded in a paper insignificant in isolation. Yet every step, however modest, is a testimony that their life matters and would be made worth living.

Contents

List of publications

Publications included in the thesis

This thesis is based on the work contained in the following papers:

I. Kroese, Adrien, Moudud Alam, Elin Hernlund, David Berthet, Lena-Mari Tamminen, Nils Fall, and Niclas Högberg. ‘3D Pose Estimation to Detect Posture Transition in Free-Stall Housed Dairy Cows’. *Journal of Dairy Science* 107, 9 pp 6878-6887 (2024). https://doi.org/10.3168/jds.2023-24427.

II. Kroese, Adrien, Niclas Högberg, Elena Diaz Vicuna, David Berthet, Nils Fall, Moudud Alam and Lena-Mari Tamminen. ‘Evaluating the automated measurement of abnormal rising and lying down behaviours in dairy cows using 3D Pose Estimation’. *Smart Agricultural Technology* 12 pp 101205 (2025). https://doi.org/10.1016/j.atech.2025.101205

III. Kroese, Adrien, Ylva Mellbin, Lena-Mari Tamminen, David Berthet, Anna Leclercq, Nils Fall, Niclas Högberg and Moudud Alam. ‘Cow cubicles and 3D pose estimation part I: how flexible head rails affect rising and lying down comfort’. *Submitted to Computers and Electronics in Agriculture on December 15th, 2025*.

All published works are available open access.

The contribution of Adrien Kroese to the papers included in this thesis was as follows:

I. Conceptualization, methodology, validation, formal analysis, investigation, writing – original draft, writing – review and editing, visualisation.

II. Conceptualization, methodology, validation, formal analysis, investigation, data curation, writing – original draft, visualisation.

III. conceptualization, methodology, formal analysis, investigation, resources, data curation, writing original draft, visualisation, funding acquisition.

Related work

Niclas Högberg, David Berthet, Moudud Alam, Per Peetz Nielsen, Lena-Mari Tamminen, Nils Fall, Adrien Kroese (2025) Exploring pose estimation as a tool for the assessment of brush use patterns in dairy cows. *Applied Animal Behaviour Science*. Volume 292, 2025, 106746, ISSN 0168-1591. https://doi.org/10.1016/j.applanim.2025.106746

In this study we evaluated different camera angles to use 2D pose estimation to detect when cows were using a mechanical brush, and which body part they were brushing. Our conclusions being (i) that pose estimation works well for that end, and (ii) that pose estimation provides continuous data which should be analysed differently than categorical data produced by human observations.

Kroese, N. Högberg, D. Berthet, L.-M. Tamminen, N. Fall, M. Alam (2024) Exploring the link between cow size and sideways lunging using 3D pose estimation. *11th European Conference on Precision Livestock Farming*, pp. 32-39
diva-portal.org/smash/record.jsf?pid=diva2%3A1916993&dswid=7796

In this study, we used the 3D pose system to generate hypotheses as to variables affecting side lunge. We regressed lunge angles on height at the withers. We found a significant effect of size on straightness with a small effect size. We also found that lunge angle was continuous with no clear breakpoint between straight and side lunge.

List of tables

List of figures

Abbreviations

AMS	Automatic Milking System
BA	Behaviour Analysis
BCS	Body condition score
CV	Computer vision
FPS	Frames per second
HD	Hypothetico-Deductive
LLM	Large Language Model
MQTT	Message Queuing Telemetry Transport
NVR	Network Video Recorder
PCA	Principal Component Analysis
PLF	Precision Livestock Farming
PoE	Power over Ethernet
SFM	Structure from motion
ToF	Time of Flight 3D imaging
WQ	Welfare Quality assessment framework
LTS	Lying to Standing
STL	Standing to Lying

Ethical statement

Animal welfare, as a field of study, is inherently tied to the management of animals under human care. While non-managed wild animals, as sentient beings, also experience positive and negative physical and mental states, these unfold mostly without human intervention. These experiences are framed within the discourse of welfare when impacted by human activities. To engage in animal welfare research is, in effect, to accept that animals will be kept in conditions shaped by human needs—such as dairy cows reared for affordable food. Welfare takes responsibility for evaluating the quality of those conditions and the outcomes they generate. The aim is not to find whether the compromise implicit in keeping animals in environments that differ from their natural state is a fair one (such as protection from freezing temperatures against restriction of movement), but rather to identify which practices best support their physical health and mental well-being, and to learn to interpret the signals animals provide about how they are faring.

1. Introduction

Assessing the welfare of dairy cows has traditionally relied on visual observations (Linstädt et al., 2024; Maroto Molina et al., 2020), whether done systematically or routinely – when a caretaker goes around the barn to check on the well-being of their animals. Observers detect subtle behaviour cues, and their ability to interpret them in context is unmatched by sensors (Smith & Pinter-Wollman, 2021). However, observers are prone to fatigue, to variable biases, and will have difficulties tracking every animal in larger herds (Hansson & Lagerkvist, 2015). As a result, at-risk animals may be overlooked until welfare issues are severe.

Dairy barns are getting increasingly digital, with sensors meant to help farmers make quantitatively informed decisions, produce more with less, take better care of their animals, and retain a feeling of control and visibility over ever larger herds. Sensor technology holds a promise to monitor all animals with the same level scrutiny and reveal individual variations, for instance in their personality (Woodrum Setser et al., 2024).

These two observations – difficulty of keeping an eye on the welfare of large herds and increased digitalization – together represented an opportunity: to use or develop technology for the purpose of monitoring welfare. Welfare is in fact a term already frequently found in publications supporting the development of precision dairy technology. If we scrutinize the specific applications of these technologies however, we notice that the parameters monitored are overwhelmingly restricted to dimensions relevant to economically efficient production. The well-being and behavioural dimensions of welfare are underrepresented (Liu et al., 2023; Stygar et al., 2021). This formed a first research gap; digital welfare monitoring.

Welfare remains a broad and complex concept, spanning physical and mental well-being. Without wanting to arbitrarily simplify welfare – a concern put forward by specialists when it comes to practical assessment methods (Foris et al., 2025) – operationalising its assessment meant that a specific focus needed to be chosen. Finding this focus led us to identify our second research gap: objective methods for posture transition assessment. While the lying down aspect of resting is well documented (notably thanks to sensor technology) transitions to the recumbent position are less explored, particularly with digital tools (Maroto Molina et al., 2020). These posture transitions provide information on how comfortable cows are in their environment (Lidfors, 1989). Visual assessment cannot provide measurements of spatial use, and sensor technology might address this limitation.

This thesis contributes to bridging both gaps by evaluating multi-view computer vision to monitor the individual variations in getting up and lying down movements, and use their measurements to infer the level of comfort offered by cubicles. We deployed, and refined, a multi-camera system over 12 dairy cow cubicles to capture their motions in 3D when transitioning between postures. The system uses synchronised RGB cameras and detects cows and anatomical key-points on frames. Using known intersecting lines of sight, it triangulates the 3D location of the key-points. The output is comparable to motion capture. This system is a prototype by Sony, who was a key collaborator on this project. They provided extensive help and expertise in deploying, maintaining and running the multi-camera system. This thesis is an interdisciplinary co-creation of knowledge between academia and industry on the place of 3D pose estimation in dairy cow welfare monitoring.

The thesis comprises three original studies. The first two focus on system development, methodological validation, and lessons learnt from automating posture transition assessment. The first study demonstrates reliable detection of posture transitions and serves as an early proof-of-concept. The second extends this validation to all key phases of both lying down and getting up, using 3D pose data to derive indicators of posture transitions. This analysis revealed that spatial use, hesitation, and lunge represent uncorrelated dimensions, suggesting that visual assessments focusing solely on duration overlook important aspects of comfort. The third study applies this system in an intervention trial, replacing conventional head and neck rails in cubicles with flexible straps. It looked at changes in posture transition under the improved cubicles. Together, these studies present a framework for interpreting sensor data into information useful to understanding parameters of cows' welfare. They showcase an example of how sensors can generate continuous data to support our understanding of cow comfort and assess welfare with a degree of automation.

2. Background

This thesis aims to bridge two research gaps:

1. Sensor technology is less commonly applied to monitoring the *good environment domain* by animal-based measures, and rarely applied to the *appropriate behaviours* domain of welfare (the concept of welfare domains (Mellor et al., 2020) will be explained in greater detail later).
2. Assessing comfort around getting up and lying down in cubicles can benefit from information on spatial use, yet the current methods are visual and cannot provide objective kinematic measures.

From these gaps, comes an opportunity to develop sensor technology for assessing getting up and lying down motions. This endeavour addresses the second gap, by developing a tool which can complement the assessment of getting up and lying down with objective measures of spatial use. By doing so, it also addresses the first gap by strengthening the body of work on sensors dedicated to the dimension of welfare relevant to comfort behaviour. This section presents an overview of trends in precision dairy technology and in welfare assessment and explains how I have come to these research gaps.

2.1 Part 1: Digitalisation in dairy production

The concept of "Precision Livestock Farming" (PLF) draws from the earlier field of precision agriculture (PA). PA represented a shift in how agricultural production is managed, wherein operational decisions are made based on individual variability rather than a field or farm basis (Taylor, 2023). PLF translates this concept to livestock farming and proposes that health management, feeding strategies and care be not homogeneous but tailored to each individual. The rationale being that feed efficiency, sensitivity to diseases and personality are individual traits, and that managing individuals separately can, if done adequately, lead to more input-efficient production and increased welfare. Whates (2008) defined PLF as follows:

> *Precision livestock farming can be defined as the management of livestock production using the principles and technology of process engineering.*

Process engineering being the continuous optimization of production systems, making extensive use of data generated during operations to fine tune operations. PLF also expands this rationale of individual focus into a time dimension, recommending all animals be constantly monitored and that care be adapted to changing responses. Continuous monitoring allows for timely interventions in case of disease, avoiding the worst consequences in terms of welfare and production if animals in a compromised state are not detected early enough. Berckmans (2017) offers the following definition:

> *The aim of PLF is to manage individual animals by continuous real-time monitoring of health, welfare, production/reproduction, and environmental impact.*

Continuous monitoring is made possible with a range of sensors. PLF has become synonymous with sensor technology and data processing for production animals. It remains important to make the distinction between the concept (real time care at the individual level) and the means (sensors to gather the necessary information at individual level).

Dairy production in post-industrial countries like Sweden has become increasingly digitalized. This ongoing trend is driven by a demand for automation, aimed at managing larger herds flexibly. It is also driven by a demand for data, to enhance visibility on herds and take quantitatively informed management decisions with a low lag-time. Some of the earliest commercially available sensors detected location using RFID (Rutten et al., 2013), and oestrus based on changes in activity (Tangorra et al., 2024). Later, computer vision has garnered a lot of interest, notably because it is none invasive and highly versatile (Fernandes et al., 2020). Demand and marketing around these technologies are often driven by concerns of competitiveness and workload manageability. A question that has emerged from both academic and industrial spheres (investigated by Stygar et al. (2021)) can be stated as follows:

"Can technical innovations in dairy cow monitoring be used for welfare assessment?".

Växa, a partner of the project supporting this thesis, conducts an on-farm welfare assessment with a framework called *Fråga Kon* (*lit.* Ask the Cow). This framework is intended as a lightweight animal-based method to evaluate welfare with animal-based measures. Acknowledging the limitations of visual observations in terms of consistency (Bokkers et al., 2012), scale (Sapkota et al., 2022) and cost (Linstädt et al., 2024), we saw an opportunity to automate this assessment. The question of which technology to use for this purpose was put forward.

It is important to state that the present thesis is not an attempt at automating a specific welfare assessment scheme, and that it is not driven by a request from the industry. Rather, it brings together academics, extensionist and industry in seeking a response element about the feasibility of automating welfare assessment.

2.1.1 Trends and gaps in PLF technology and the data it generates

PLF has gained significant momentum over the past decades and is currently seeing its highest publication growth rate (Marino et al., 2023). Dairy-cattle monitoring has shifted from single-purpose sensors (e.g., typically pedometers for oestrus detection, and milk electrical conductivity) towards integrated, systems that continuously analyse data on behaviour, physiology, production, and location to recommend actions (like insemination or treatment). Wearable accelerometers, usually placed on leg, collars or ear form the majority of commercially available monitoring solutions (Stygar et al., 2021), supporting heat detection (e.g. Cattle Watch, South Africa), activity/rest (e.g Cow Manager, Sweden), feeding/rumination proxies (e.g. Real Time, Boumatic, USA), and time-budget metrics (e.g. MooMonitor, Dairy Master, Ireland).

A technology that has gained significant attention in recent years for its flexibility and non-invasive nature is computer vision (CV). It commonly relies on automated analysis of 2D video in the RGB spectrum. Computer vision benefits from high versatility; a review on the potential to automate WQ presented mainly computer vision as a potential technology to measure indicators that did not yet have a dedicated sensor (Maroto Molina et al., 2020). 3D Computer vision has also seen substantial developments. Several different techniques have been applied to obtaining visual information on cows in 3D. One method is the time of flight (ToF), which translates the time it takes for near infrared light to be reflected onto the sensor into a distance. This was applied to estimating cow body weight (Jang et al., 2020). A way of generating similar 3D images is stereo active infrared (IR), which projects an IR pattern onto a scene and analyses differences in the pattern between two closely located sensors to infer volume. This method was applied to determining cow postures with the goal of identifying anomalies (Lee et al.,

2024). Finally, another method for 3D computer vision is to triangulate keypoints from 2D pose estimation between several cameras (Huang & Moliner, 2022; Kroese et al., 2024). 3D pose produces only a limited number of points, whereas ToF creates a point cloud with volumetric information. It does have the advantage of needing only affordable 2D cameras. Pose estimation has applications for monitoring certain behaviours, for example mechanical brush use (Högberg et al., 2025). It can potentially substitute motion capture for measuring kinematics in situations where the latter is too impractical (Lawin et al., 2023), such as movement amplitude when getting up and lying down (Kroese et al., 2025) which will be the main focus of this thesis.
When reading works on precision dairy, a pattern seems to emerge: the focus is mainly on health, nutrition and reproduction (Palczynski, 2019). This pattern was found in reviews of the existing literature (Liu et al., 2023; Supplementary material by Stygar et al., 2021). Rutten et al. (2013) document the typical scope of PLF technologies; they note that, as of publication, most sensors were dedicated to mastitis (25%), fertility (33%), locomotion (30%) and metabolic disorders (15%). They attribute their predominance to economic importance but also to the maturity of their fields of research.

I wanted to get a quantitative insight into the themes that fellow academics in precision-dairy had explored. I ran a bibliometrics analysis of original research papers with the keywords “precision livestock farming” and either “dairy”, “cow” or “cattle” in the Scopus database. I identified 271 studies with indexed keywords. I retrieved the keywords used for each study and produced an exhaustive list of unique keywords. I asked Mistral (a LLM) to group the keywords into themes, and it identified the following: nutrition, heat stress, productivity, body condition, reproduction, diseases, behaviour, welfare and environment. I asked it to make disease into a broader health theme, to add a management theme, and to narrow down environment to housing (including pasture for outdoor animals). I then asked it to group all studies in one or more categories. My aim was to see how frequently welfare occurred and what themes it co-occurred with.

Welfare was a common theme, appearing as a keyword in 85 out of the 271 studies (or 31%). “Welfare” was second to “Behaviour”, which appeared in 41% of studies. If we look at the co-occurrence of welfare with other themes we get the following:

Table 1. Co-occurrence of welfare with other themes from bibliometric analysis of keywords in precision dairy studies.

Themes co-occurring with welfare	Number of studies	Frequency in original selection
Behaviour	41	15
Health	32	12
Heat stress	32	12
Housing	32	12
Productivity	16	6
Management	14	5
Body condition	11	4
Nutrition	10	4
None	10	4
Reproduction	5	2

We do see in Table 1 that behaviour is the top theme co-occurring with Welfare. The predominance of behaviour is expected, since behaviour serves as a proxy to derive other information, such as compromised health (Högberg et al., 2019), feeding (Riaboff et al., 2022) or mental states (Keeling et al., 2021). In fact, it was mainly stated in association with the themes of health, housing and heat stress. 21 studies referred to welfare or behaviour without a health, fertility or productivity component (it would be a stretch to claim that they address the behavioural aspect of welfare based only on the keywords). This contrasts with 83 studies having a health component.

Welfare seems to be a ubiquitous theme, perhaps used as an umbrella term for any sensor technology which has the potential to assess the state of the animal or even improve the life of animals (or farmers). Welfare is predominantly mentioned in the context of health, or other parameters affecting management, such as reproduction or heat stress. Its behavioural dimension with other applications than health or management is rarer.

2.1.2 Research gap 1: sensors dedicated to welfare assessment

Upon reviewing drivers of precision livestock farming and available technology, an imbalance between welfare and production parameters emerges. According to a bibliometric analysis of themes in PLF publications (not mine this time), the keyword “health” is the earliest (2014) to become prominent in scientific articles focused on technology for dairy farming (De

Oliveira et al., 2024). The keyword "estrus detection" reaches a comparable prominence in PLF publications in 2016 and "body condition score" in 2018. "Welfare" appears later as a trending keyword associated with PLF in 2019 (De Oliveira et al., 2024). While health of the animals is an important domain of animal welfare (and of producers' mental health) – and while production can be an indicator of good health – producers, researchers and consumers recognize that animal welfare extends beyond physical well-being (Skarstad et al., 2007; Alonso et al., 2020).

Stygar et al. (2021) identified 30 validated sensors aimed at welfare monitoring presented in scientific articles and 129 additional retailed technologies, 18 of which were externally validated. They found varying results regarding performance in classifying behaviours and inferring welfare states, but performance is not the point at hand. Most tools had applications for good health and feeding dimensions of welfare. Wearable accelerometers had potential to assess good housing, but it was not always their purpose. They conclude that available PLF technologies currently have low potential to assess behavioural indicators of welfare. If we further scrutinize the full list of commercial technologies in Stygar et al.'s 2021 supplementary materials, we see that only 5 tools incorporated indicators of resting behaviour and 1 was primarily focused on welfare parameters other than health, reproduction, feeding and locomotion. If we look into new entries since this review was published, we note the release of the DeLaval Plus Behaviour Analysis (DeLaval international, Sweden) which notably monitors time budget in resting area. We also have the TriAct package to detect abnormal motions in cows' resting behaviours and posture transition movements using accelerometers (Simmler & Brouwers, 2024).

A review by Liu et al. (2023) identifies the following applications of PLF in dairy: individual recognition, behaviour monitoring, disease detection, BCS and feeding. Within the 23 publications fitting the *behaviour* application, 5 studies dealt with lying (and standing) behaviour, 9 with feeding behaviour (including rumination), 6 with oestrus and 4 were concerned with classifying activity type. For the 5 studies focused on lying behaviour, their rationales for automating behaviour monitoring are the following:

- Detecting deviations in lying behaviour as an indicator of compromised welfare (Achour et al., 2019)

- Technological advancement (Balasso et al., 2021; Tamura et al., 2019)
- Detection of idle time (Ma et al., 2022)
- Monitor indicators of economic performance (Balasso et al., 2021; Wei et al., 2023)

The term "welfare" is mentioned 4 times as a technology application by studies reviewed by Liu et al (2023): once in the context of lameness, twice in the context of BCS and once regarding oestrus behaviours. In two of the mentions, general health is stated as a motive for assessing welfare.

Because of this discrepancy between health-related welfare parameters and welfare as a broader concept, technology developed with the main purpose of guiding health and production management likely has blind spots when it comes to assessing welfare. A notable work supporting that is by Barry (2025) who evaluated the possibility of monitoring dairy cow welfare using data available through National Milk Records (Original Norwegian: Kukontrollen). They compared the routine data to Welfare Quality and did not find the records adapted to evaluating welfare in its multi-dimensional aspect.

2.2 Part 2: welfare assessment of dairy cows

What is on-farm welfare assessment? Is it a veterinary inspection? Is it reporting for compliance with regulation and certification schemes? Is it a farmer checking if their animals are doing well? Is it a systematic assessment of the state of the individuals according to strictly defined indicators? In this context, we will accept all these possibilities and define welfare assessment as "checking if an animal is doing well". What the term "well" implies will be discussed below. We will however operationalise welfare assessment through the lens of the last aspect: an assessment of key indicators on the animals, chosen by rigorous validation, such as is done in Welfare Quality (Blokhuis et al., 2013).

2.2.1 Understanding of animal welfare

Animal welfare is a complex multi-dimensional concept, encompassing both physical and mental well-being. Animal welfare science has evolved from an interpretation of welfare focused on the avoidance of negative experiences (negative welfare) to an increasing recognition that welfare must integrate the presence of positive experiences (positive welfare) (Rault et al., 2025). A framework that includes both positive and negative dimensions is the 5 domains of animal welfare: good physical health, nutrition, and environment, positive mental states, and appropriate behaviours (Mellor et al., 2020). For this thesis, we will use this definition of welfare. The motive is firstly that the 5 domains are broad enough to encompass both the notions of freedom from negative experiences, of positive states, and of behavioural agency. We note that freedom from negative experiences does not mean complete absence of them, but rather a perceived ability to cope with them (Broom, 1996) and a balance of experiences remaining positive towards what can be summarised as "a life worth living" (Mellor, 2016). Secondly, this definition of welfare can be operationalised for welfare assessment; the well-established Welfare Quality Protocol uses these same domains, with the difference that it groups "positive emotional states" under "Appropriate Behaviour" (Blokhuis et al., 2013). Through the lens of the 5 domains, we can evaluate both the animals (good health and positive mental states), their housing (good environment) and the keepers' practices (notably good nutrition but not restricted to it).

2.2.2 Cubicle systems in larger dairy operations and their implications for welfare

Average dairy herd size has substantially increased over recent decades, in Europe and globally (Barkema et al., 2015). In Sweden, the average herd size grew from 34 cows in 2000 to 102 cows in 2021 (DG Agriculture and Rural Development, 2021). This trend is part of a broader pattern observed throughout Europe, where average herd size varies widely by country and region but has increased in response to economic pressures and industry intensification.

The implications of growing herd size for animal care have been the subject of research and debate (Barkema et al., 2015) but the link between herd size and welfare is not trivial. Evidence suggests higher welfare in larger operations, benefitting from greater professionalization, standardized routines, and technologies that support monitoring and animal health (Beggs et al., 2019; Lindena & Hess, 2022). This effect of herd size is however small compared to the variability between farms. Each system has its compromises: smaller herds are more likely to use tie stalls restricting movement, while larger herds tend toward zero-grazing (Barkema et al., 2015; Legrand et al., 2009). Although advanced tools in large herds can improve detection and care, effective action still depends on farmer training and engagement. Ultimately, as herds grow, the concern is not that of level of care, but whether individuals with compromised welfare risk getting overlooked (Hansson & Lagerkvist, 2015).

The most common housing in large-scale intensive systems is the free stall, where cows can move freely. The stall is designed in such a way that space-use efficiency is maximised, and the risks contamination minimized. Cubicles are designed in such a way that cows lie down with their rear over the alley to avoid soiling the bed (Gieseke et al., 2020). This is notably achieved by an intentionally restrictive neck-rail. It is placed in such a way to prevent a cow from extending too far into the cubicle when initiating the lying down motion, and to encourage lying over standing. Cubicles will often have either a head rail or a brisket board to act as a frontal limit. The result is that, while head and neck rails keep beds mostly clean, cubicles have effects on comfort around posture transition (Lidfors, 1989). A review by Nielsen et al. (2023) emphasizes the detrimental effects of inappropriate cubicle on dairy cow welfare, including hock lesions, claw disorders, and increase lameness prevalence, alongside a potential impact on mental states. Cows in restrictive cubicles will take longer to rise and lie down (Brouwers et al., 2024), will display shorter movement amplitude (Ceballos et al., 2004) and more frequent of abnormal motions (Brouwers et al., 2024). Abnormal motions that cows display in cubicles include side lunge (Brouwers et al., 2023a,b), abnormal order of motions (rising front first) (Lidfors, 1989) or hind-quarter stepping (Zambelis et al., 2019).

2.2.3 Welfare assessment methods around cubicle comfort

On-farm welfare assessment is structured around validated frameworks; practical protocols using a variety of indicators. Welfare Quality (WQ) (Welfare Quality® Consortium, 2009) is one such notable framework, which is validated (in the sense that the prescribed observations correlate well with the state of the entire herd) and used as a gold standard in welfare assessment of production animals (Linstädt et al., 2024). Assessment is performed by trained assessors, using mostly direct observations of the animals but also records and evaluation of the environment. WQ offers a standardized protocol leading to highly comparable results across contexts

In its evaluation of housing, WQ assesses notably the cubicles with two indicators: the number of cows taking more than 6.3s to lie down, and the occurrence of collisions in the process. The time needed to lie down denotes hesitation, or intentionally slow movements to avoid painful collisions.

The demands of WQ (in terms of time) have prompted the adoption of lighter protocol, such as the Swedish *Fråga Kon* (Växa, Sweden) or the Danish Dairy Cattle Federation's (DCF) protocol that instead look at the time needed for a cow to get up. In a later report, the European Food Safety Agency (EFSA) identifies risks to the welfare of production animals and offers metrics to quantify theses risks. To evaluate movement restrictions and resting problems, they propose gait, hygiene, lesions and deviations from normal lying down and rising up movement (Nielsen et al., 2023). In other assessments of stall comfort, lying time and bout frequency are predominantly used as measures of comfortable cubicles (Cook et al., 2005).

2.2.4 Research gap 2: continuous spatial use measurements in posture transition assessment

Posture transitions were chosen as a focus for this thesis. As we have seen, they are relevant to comfort and welfare and are used in welfare assessment accordingly. Yet, their assessment is mostly summarised by the time dimension, which does not offer a full picture of the behaviour, and objective methods to obtain a more fine-grained picture are lacking.

A study aimed at evaluating the potential for automating Welfare Quality found that overall, many sensors or their combination could be used, or are already used, to measure most of the criteria from WQ (Maroto Molina et al., 2020). They do however state that *"No references to sensor systems enabling the measurement of time needed to lie down, collisions with equipment or cow positioning in the resting area were found"*. This quote omits motion capture, which can do just that (Ceballos et al., 2004), but the latter technique is not practical in barns, thus limited to specific trials, and not large-scale monitoring.

Relevance of posture transitions to welfare

Posture transitions are biologically essential activities: cows must rise and lie down repeatedly in order to access feed, water, and rest. Cows spend more than half of the day lying (Munksgaard et al., 2005; Tucker et al., 2021; Wegner & Ternman, 2023) and alterations in lying time and bout structure have been associated with lameness (Ito et al., 2010; Thompson et al., 2019), stall design (Brouwers et al., 2024), and health status (Von Keyserlingk et al., 2009).

Posture transitions are physically demanding (Schnitzer, 1971). Difficulty or reluctance in performing these transition motions can reflect pain, discomfort (Lidfors, 1989) or inadequate space allowance (Cook, 2009). The frequency of posture transitions seems to be affected by the comfort level of the bed, suggesting that cows are more reluctant to display these movements in an unsuitable environment (Haley et al., 2001). The EFSA identifies bad housing as a risk to welfare, notably because it hinders proper posture transitions (Nielsen et al., 2023). Welfare Quality recognises that unhindered, swift posture transitions are part of cows' opportunities to display natural behaviours and reflect comfort (Blokhuis et al., 2013). For example, restrictive or poorly designed cubicles can force cows to alter the trajectory of rising, increasing collision risk and delaying movement (Cook & Nordlund, 2009). Bedding material affects both the ease and willingness to lie down; cows provided with deep-bedded sand stalls exhibit shorter lying-down durations and more frequent bouts compared to those on mattresses (Haley et al., 2001). Reduced number of posture transitions may therefore serve as an indicator reflecting an environment where either the recumbent position is uncomfortable or the act of getting up and lying down is difficult.

Existing methods and their limitations

Earlier work on posture transitions has highlighted the importance of space allowance, particularly the ability for cows to lunge their head forward when getting up (Cook, 2009). A kinematic comparison of cows lying down showed that they used less space in cubicles than in open packs (Ceballos et al., 2004). The effect of insufficient space can be seen on cows, for example with larger cows taking more time to get up, and having more hindquarter readjustments when lying down (Zambelis et al., 2019).

Clues of uncomfortable cubicles can be seen on the cow, such as neck and dorsal lesions occurring from contact with the environment when getting up or lying down (Zambelis et al., 2019). These clues can also be seen on the cubicle bars, specifically if they are polished in specific spots like under the head rail, meaning that cows regularly come into contact with it. However, studies of cubicle comfort largely employ cow comfort index (proportion of occupied cubicles with cows lying down) (Cook et al., 2005), lying time

(Abade et al., 2015; Haley et al., 2000), and preference (Abade et al., 2015; Tucker et al., 2006).

Direct visual observation has traditionally been the main approach to assess posture transitions, yet this method comes with several important limitations. First, observations are constrained to the moment in which the behaviour occurs: once a cow has stood up or lain down the opportunity to assess details of the transition is lost, as the observer cannot rewind live events. Second, visual observation does not allow for quantitative measurement of spatial use or fine-grained kinematic detail. Parameters such as displacement trajectories, joint angles, or timing of limb movements are better studied with sensors, motion capture being a gold standard (Lawin et al., 2023). Third, visual scoring is time-consuming and resource intensive. Trained observers are required, and assessments are performed irregularly. The Welfare Quality protocol, for example, requires approximately one full day of observation per farm (Linstädt et al., 2024). It is impractical for high-frequency monitoring. Lastly, observer bias also remains a possible limitation, though training and calibration can reduce variability; Zambelis et al. (2019) demonstrated very high inter-observer agreement (k=0.93) in scoring abnormal lying-down behaviour. While visual observations provide valuable qualitative insight, their inability to deliver continuous, objective, and quantitative spatial-use measurements limits its application for sensitive and scalable welfare monitoring of posture transitions.

Sensor systems present an opportunity to automate the assessment of selected welfare indicators (Maroto Molina et al., 2020), not the least posture transitions (Brouwers et al., 2023b). We have already mentioned the use of motion capture, which offers fine-grained kinematic information at the expense of practicality.

Brouwers et al. (2023b) sought to detect abnormal rising and lying down movements using accelerometers and supervised learning. Accelerometers are well adapted to production settings. Their work lead to the creation of an R package for analysing rising and lying down movements (Simmler & Brouwers, 2024). When Brouwers et al. (2023b) attempted to automate the detection of sideways lunge using accelerometers, they reached moderate accuracy (65%). This is encouraging in terms of technical developments but insufficient for practical use. The authors impute this to a discrepancy between the way data was labelled (straight vs angled lunge) and the continuous nature of sensor data. There were many misclassifications on ambiguous edge cases. The author of the aforementioned study later suggested that “ethograms should be machine-learnable” (also noted by Gris et al. in 2017). The work presented in this thesis draws on this discovery and will explore how sensor data can be interpreted in novel ways beyond trying to automate existing measures.

2.3 Bridging two gaps with digital methods for welfare assessment

On one hand, sensor technologies for health monitoring have advanced significantly, but there remains a gap in methods for assessing how comfort influences behaviour. On the other hand, posture transitions—essential for expressing lying-down comfort behaviour—are assessed using limited visual and categorical methods that would benefit from more objective approaches. There is thus an opportunity to address both gaps by providing a sensor-based method for posture transition assessment.

This project exemplifies the kind of integrative approach advocated by Foris et al. (2025) where expertise from engineering and ethology are joined to co-create solutions that are technically sound and relevant to improving welfare. Automated evaluation of posture transitions using 3D pose estimation requires technical expertise, to design robust algorithms capable of capturing subtle and rapid movements, and to deploy them on the appropriate suite of hardware. Animal welfare expertise ensures that the indicators derived are meaningful within the biological and behavioural context of dairy cows.

Posture transitions are well suited for this co-creative framework: first, they are discrete, repeated events that can be quantified objectively by computer vision thus scaled and compared. Secondly, they carry biological significance as indicators of comfort, health (Lidfors, 1989), and the suitability of housing systems (Cook & Nordlund, 2009). Thirdly and finally, applying this technology goes beyond telling us *“what is wrong about the cow”* but generates animal-based evidence on the suitability of housing systems that can be used for their improvement (Brouwers et al., 2024).

2.3.1 Pose estimation in 3D

Pose estimation in 3D is a field of application of computer vision technology aimed at predicting the (x,y,z) coordinates of keypoints (typically joints and other body parts) from 2D images. There are several lines of development. One is to triangulate synchronized detections across several 2D images (Huang & Moliner, 2022). Another other is to lift 2D into 3D using neural networks trained on ground truth 3D coordinates (Gosztolai et al., 2021). A final method is to predict key point location directly on depth images (Ye et al., 2011).

In this thesis, we worked with the first technique: 3D pose from multi-view fusion of 2D poses. This method will be the sole focus. The following section will provide a high-level overview of the method, including computer vision, pose estimation in 2D and multi-view fusion.

Pose estimation

Pose estimation is a task of computer vision where the aim is to locate the coordinates of key-points. A subject is detected on a frame as a set of points (joints) and linkages (bones). Detecting the keypoints usually relies on the combination of two techniques: neural networks to detect the location of the points, and geometric constraints to refine the pose based on plausible linkage length and joint angles (Nogueira et al., 2025).

3D fusion of pose estimation

The process begins by capturing synchronized images from several cameras positioned around the subjects. The cameras need to be intrinsically calibrated, that is, determining intrinsic parameters to align the camera's coordinate system with world coordinates (Moliner et al., 2021). Pose estimation in 2D is run independently on the frame from each camera. At this stage, the algorithm used for 2D pose has little relevance but accurate detection of the key-point from all angles is conditional for precise 3D estimation. The 2D keypoints from the different views are geometrically combined using known intersecting lines of sight. An overview of the specific process for determining these lines of sight will be presented further in the methods section. An example of pose estimation in 2D can be seen on the upper frames on Figure 1.Figure 1. Pose estimation in 2D and 3D fusion of two cows. The blue cow is in the lunge stage of rising. The result of 3D fusion is shown as stick figures in the lower part.

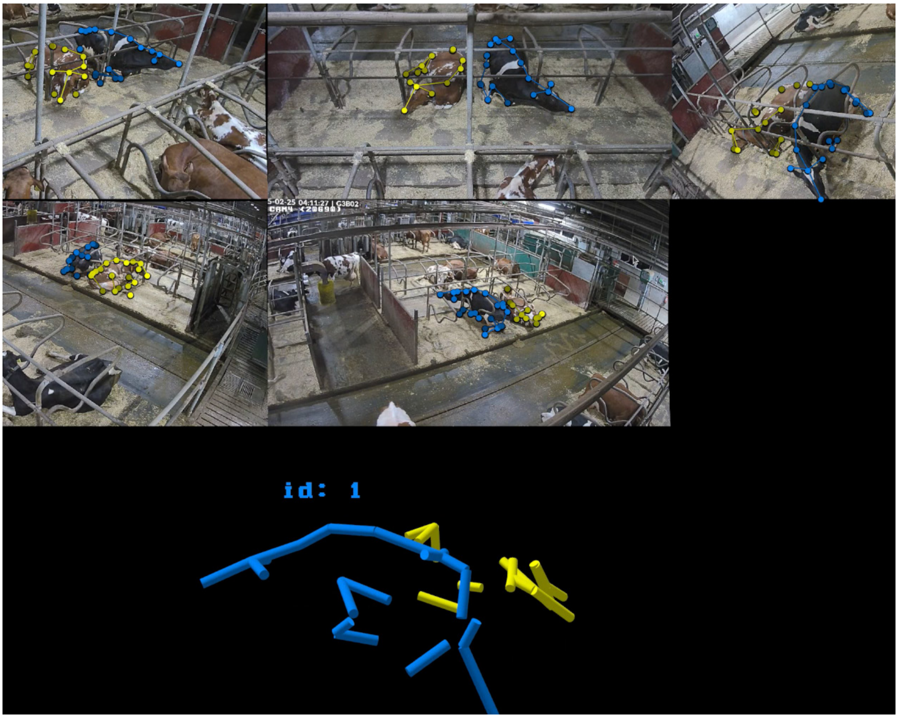


Figure 1. Pose estimation in 2D and 3D fusion of two cows. The blue cow is in the lunge stage of rising.

Direct triangulation can suffer from noise, occlusions, and inaccuracies in 2D detections, which degrade 3D pose quality. To improve robustness, algorithms incorporate confidence weighting, optimization based on reprojection errors, and kinematic constraints such as limb length, joint order or symmetry to refine initial 3D estimates (Moliner et al., 2021). Reprojection error minimization refines 3D joint positions by iteratively adjusting. Each iteration seeks to minimize the discrepancies between the original 2D keypoint detected on a camera's frame, and the reprojection of the 3D estimate of that keypoint against that same camera's line of sight.

Pose estimation has various applications across scientific research and practice where the change in position of body joints is of interest, for example gait of horses (Lawin et al., 2023), neural responses of animals (Gosztolai et al., 2021), or motions of athletes (Qu et al., 2024), and as we will see, cow comfort.

2.3.2 Posture transitions of dairy cows

Rising

Rising movements follow an innate sequence of motions, successively soliciting specific body regions and muscles groups. Typically, a cow begins from a lying position by extending her neck and head upward, which shifts the centre of gravity forward and prepares for limb engagement. Next, the forelimbs are folded inward, resting on the carpal joints, with the withers rising slightly as a result. Cows are in some cases observed to crawl backwards at this stage, hypothesised to be an attempt to increase the space available in front of them. Crawling is an important qualitative criterion of insufficient space used in the *Fråga Kon* framework. This behaviour is not observed in open environments. The cow then lunges her head forwards. The forward lunge movement displaces the weight of the cow forward, away from the hind legs and onto the front limbs (Schnitzer, 1971), with approximately 2/3 of the total weight being born on the carpal joint at this point (von Metzner, 1978). Moving weight into the front allows the cow to lift its hind legs and position them under the body in a swift motion. The cow then moves upward, pushing with the forelimbs while shifting weight onto the hind limbs, which extend to elevate the pelvis and sacrum. This action is followed by final straightening of the forelimbs until the cow has assumed an upright position. Cows are often observed to stretch their back in an arch shape after getting up (Schnitzer, 1971). A visual of the sequence of movements is illustrated in Figure 2. As a reader you might want to bookmark this figure since it will be referred to quite often.

Cubicles are designed to allow cows to get up and down while prioritising efficient spatial use and cleanliness. A Trade-off exists between hygiene and comfort (Bernardi et al., 2009). Studies and guidelines vary in their recommendations regarding design and dimensions of cubicles. Lunge space is particularly important to the rising motion (Cook, 2009). Cows require clear, unobstructed forward and lateral space to complete rising motions without restriction or injury. Increasing lunge space can effectively decrease the frequency of abnormal motions, highlighting the importance of stall and cubicle design for improving welfare (Dirksen et al., 2020). One study suggests at least 0.9m of forward lunge space (Cook, 2019) but elements like the head rail can substantially interfere with how much cows can make use of this forward space.

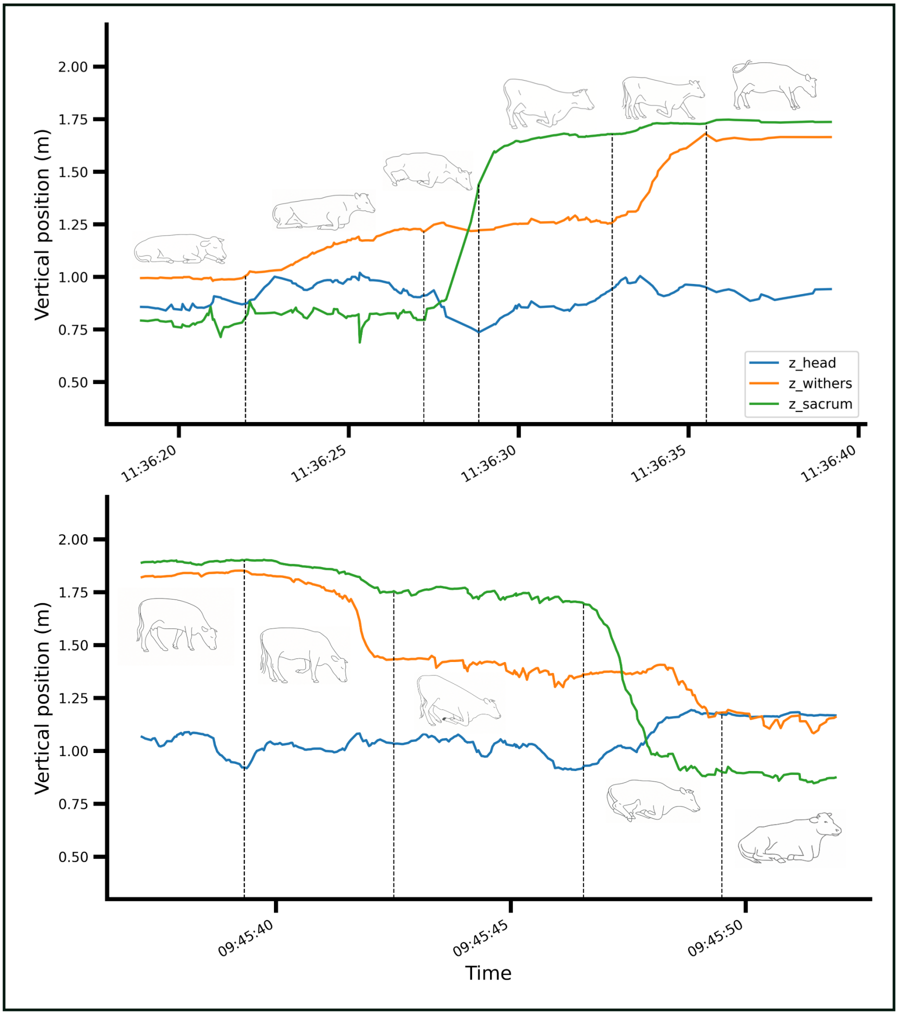


Figure 2. Characteristic vertical movement patterns of the head, withers and sacrum keypoints during rising (top) and lying down (bottom) with stages/phases marked by dashed lines.

Lying down

Before lying down, once a lying spot has been chosen, cows swipe their heads to the sides as they inspect the ground. These are referred to as intention movements (Krohn & Munksgaard, 1993; Lidfors, 1989). They bend one leg then the next and descends until it rests firmly on its carpal joints. In some cases, hind limbs are readjusted away from the side they will be resting on. The cow stretches its head forwards and down as it and the hind legs are lowered. Finally, the cow lets herself fall gently on the flank. The legs are then usually tucked under the body.

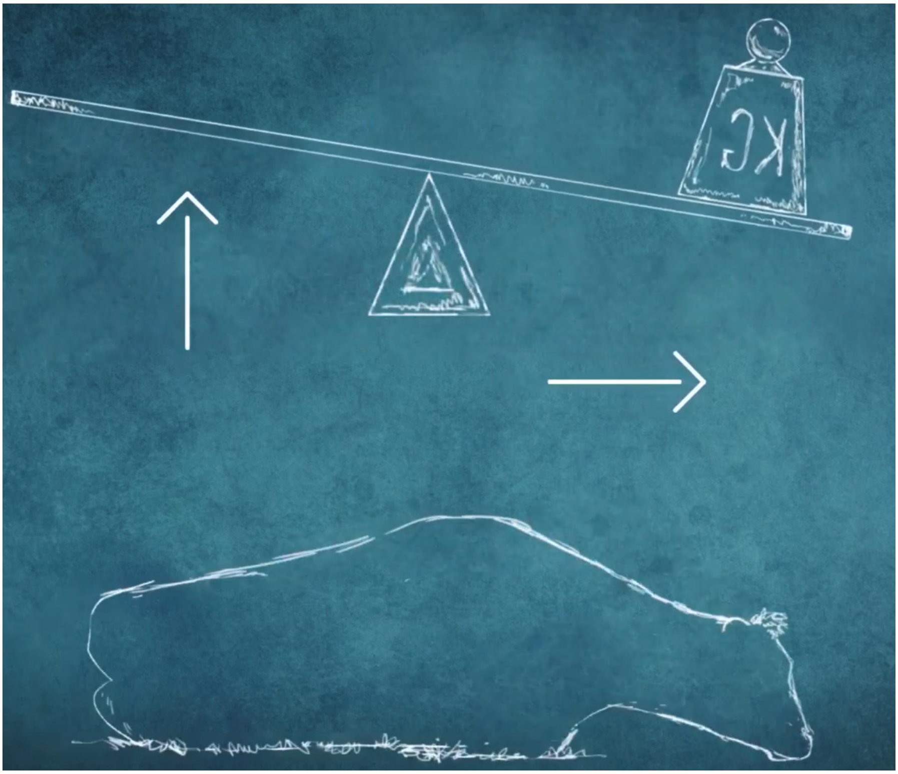


Figure 3. Excerpt from an instructional video explaining the leverage effect of the head lunge. Credit: Växa Sverige, partner of this project.

Cows can use 2.6m to 2.9m in total longitudinal space when lying down (Lidfors, 1989). They use approximately 0.7m to 1m of lateral space (120% to 180% of hip width) (Tucker et al., 2004) and up to 10.9m forward space (Ceballos et al., 2004). The largest horizontal movements of the hip typically occurred at two average heights: one between 0.9 and 1.35m (Ceballos et al., 2004; Tucker, et al., 2004). After the posture transition is complete, space will also affect lying behaviour: animals spent an additional 42 min/ 24 h lying in stalls measuring 1.26m in width compared to stalls those 1.6 cm wide (Tucker, Weary, & Fraser, 2004).

3. Aims of the thesis

There is a trend in farming systems of post-industrialized countries seeking to automate animal welfare monitoring (Barry et al., 2024; Buller et al., 2020; Maroto Molina et al., 2020). This trend is the response to both a preoccupation and an opportunity. The preoccupation is that herds get increasingly large and intensified. These intensive systems do not imply bad welfare (Lindena & Hess, 2022) but individuals in such systems with compromised welfare risk going unnoticed (because there are so many animals per caretaker). The opportunity is that digital technologies could monitor welfare parameters, objectively and continuously, providing real-time information on how well each individual cow is faring. The overall project that this thesis is a part of fits in this trend of automated welfare monitoring.

The project originally asked the question *can we automate welfare assessment?* The *can* term implies not that we meant to automated welfare assessment but is rather about exploring what is possible. *Automate* originally meant reproducing with sensors indicators that are already established, but I will talk about lessons learnt on this matter. W*elfare assessment* refers narrowly to indicators for animal-based measures which are validated and have an established link with welfare outcomes.

This thesis seeks to contribute to the body of evidence and the available technology for automated welfare monitoring. It exemplifies the automated monitoring of animal-based measures by presenting a sensor-based solution to evaluate the quality of posture transitions. The specific aims of the thesis are:

- To develop data management and processing tools to generate insights on cow comfort from continuous pose estimation data (Paper I, II & III).
- To explore a use case for multi-view computer vision addressing limitations in visual assessment of posture transition comfort (Paper II & III).
- To apply the system for the improvement of cubicles using animal measures in a commercial setting (Paper II & III).

4. Overview and comments on materials and methods

This section will explain, mostly chronologically, the steps undertaken, and the methods developed to achieve the aims. It will present a general overview of the data collection, management and processing procedures Specific and repeatable descriptions of each method used to generate results can be found in their respective papers

This thesis lies at the intersection of computer vision, biomechanics, and animal welfare science. A reader expecting to deepen their expertise in computer vision might be left unsatisfied from the only high-level overview of the technology. In the same way, a biomechanics expert might not find the level of detail they expect regarding the movement of anatomical structures. An expert in animal welfare might find the operational definition of welfare too pragmatic. This study sits at the crossroads between these three fields. Rather than diving into one of these, this thesis attempts to bridge them, explaining how data on joint kinematics was captured with computer vision with the aim of informing on welfare indicators.

Previous research has already sought to quantify the spatial use and displacement of different anatomical structures during posture transitions. A few decades ago, a grid was placed behind the cow that was recorded with film during posture transitions. Using known perspective coefficients between the camera, the cow and the grid, researchers were able to quantify the total longitudinal space used by cows, and the displacement of the head when getting up in unrestricted environments (Schnitzer, 1971). Later, motion capture was used to measure again total longitudinal space and head displacement with the addition of lateral space used this time when lying down, using motion capture (Ceballos et al., 2004). Motion capture allowed precise comparison of spatial use and movement patterns in cubicles versus open packs. These techniques generated important insights in spatial use requirement of cows that are used to inform cubicle design and to derive indicators of abnormal posture transitions. These techniques do have the downside of requiring a controlled environment. This limitation might explain the low sample size (n = 5 cows) in the latter study.

Pose estimation in 3D from multi-view computer vision acts as a form of markerless motion capture. It may not offer the same level of stability and granularity as true motion capture with reflectors but has the practical advantage of being scalable without intervention on the animals. This enabled us to monitor the movements of 183 and then 85 cows in what I will refer to as "production settings". By production settings, I mean conditions regarding the physical environment, diet and daily activity patterns of cows such that would likely be found on commercial farms. The research was conducted in the dairy barn of the Swedish Livestock Research Centre, which in fact works as a commercial farm with the addition of research and education activities. For this thesis, we deployed a multi-camera system to learn to adapt it to deliver actionable insights on the animals' welfare. The system outputs the coordinates of selected body parts in a 3D space. This data needs to be further interpreted into information on the animal.

4.1 Animals, housing and timeline

The multi-camera system for 3D key-point acquisition was set-up in one of 5 dairy pens of the Livestock Research Centre in Uppsala, Sweden. Each pen had a milking robot with voluntary access The cameras covered 12 out 66 cubicles. The cubicles model is C1300 (DeLaval, Sweden) from 2010. They consist of a 2.1m by 1.25m lying surface with neck rails. The beds were covered in peat in Phase I and with straw in Phase II. Bedding was replenished several times a day by an automatic dispenser. The end of the cubicle is marked by a head rail and a concrete step. Rows of cubicles are facing each other with 1.65m between the front end of each row. Dimensions and design of cubicles can be seen on Figure 4. The floorplan of the pen and the covered cubicles can be seen on Figure 5.

Except during the experiment (March – April 2025), cow traffic in the pen was independent of data collection and was based on the management needs of the farm. A first data collection phase (phase 1) ran from December 8th 2021 to April 28th 2022, a total of 183 cows were present in the research pen, although lack of individual identification prevents us from knowing exactly which individuals expressed the recorded bouts. Sequences showing bouts were extracted for 32 of those days This material was used for papers 1 and 2. A second phase of data collection (phase II) ran from February 24th to April 7th, 2025. This phase corresponds to an experiment, where the head and neck rails were replaced by flexible straps, 85 different individuals were present in the pen. More details on that experiment will follow.

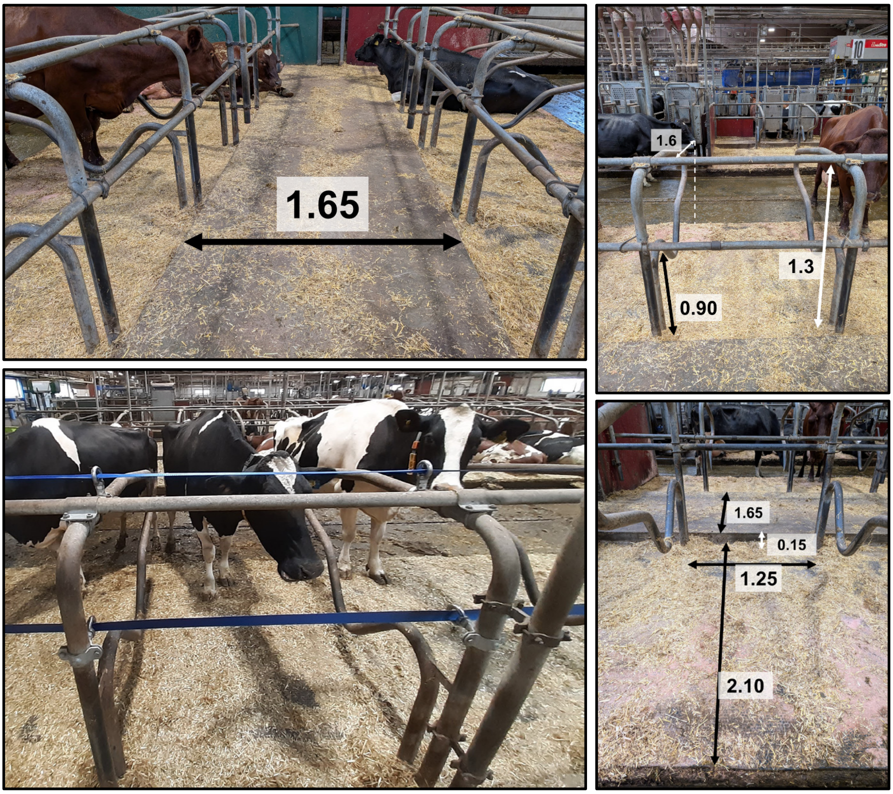


Figure 4. Cubicle design and dimensions (meters). Flexible head and neck straps in the lower left quadrant.

4.2 Multi-camera setup and data acquisition

The multi-camera system is a proprietary technology of Sony (Sweden). Despite its novelty and technological prowess, the method for acquiring 3D pose will not be the focus of this thesis, for the simple reason that it is not my own work. I will still present an overview for the sake of understanding.

The data used for this study are the synchronized video recordings, and the poses 3D generated from the video. Let's establish some terminology. A **scene** is defined as the number, placement and posture of cows in the area covered by the cameras at a given timepoint, which is captured on several synchronized frames and for which there is position information. Each camera has a **view** of the scene, meaning its orientation and field of view. A **frame** is the tensor of pixels produced by a camera at a given timestamp. The frames for one scene are synchronized meaning that they share the same timestamp. The system is robust against synchronization misalignments of up to 0.5 seconds for motions equivalent to human walking (Huang & Moliner, 2022). In terms of computer vision outputs, an **object** corresponds to the detection of a single cow. In 3D, each to each object corresponds a set of **keypoints** which are the (x, y, z) coordinates of specific anatomical landmarks. The absolute coordinates are irrelevant, but their change is expressed in meters and informs on motion amplitude. The keypoints pertaining to an object together form a **pose**, which is a set of points and linkages describing the location of the anatomical structures in space from which we can derive posture of each animal. The poses corresponding to the same object across successive scenes are identified by an id, which will be referred to as **track** number. In the text, I will often refer to a **sequence**, which is the snippet (10 to 60s) of successive poses centred on a posture transition, along with the video from all cameras used to generate it

4.2.1 Physical installation

3D fusion of pose estimation from multi-view computer vision, as its name suggests, requires at least two cameras to provide a 3D pose (R. Hartley & Zisserman, 2003) but more cameras increase robustness of the triangulation and reduce sensitivity to occlusion. There is no theoretical maximum number of cameras, this is more a concern of cost, practicality and processing capacity. The first phase of data collection was done with 6 then 7 cameras (G3Bullet, Ubiquiti, USA). Then, the experiment used 9 cameras organised in 2 groups of 5 and 4 calibrated separately.

Cameras

The cameras covered 12 cubicles. A different set of cubicles was used for development (phase I or papers I and II) and the experiment (phase II or paper III) with 4 cubicles overlapping between the phases. The cameras were positioned between 1.8 and 3.6m high, oriented towards the cubicles so that all points in the cubicles would be visible by at least two cameras at all times.

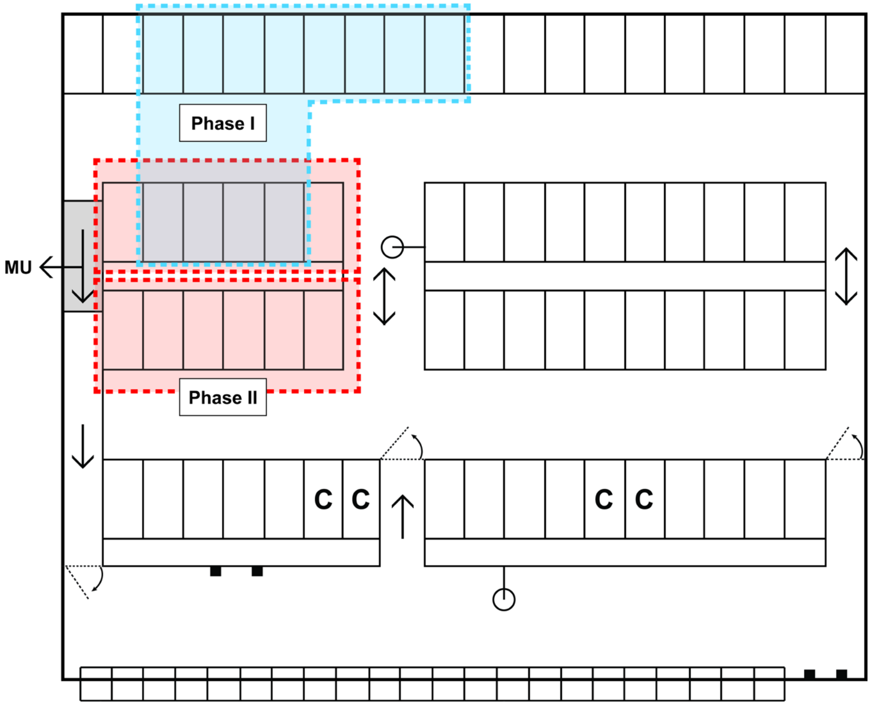


Figure 5. Floorplan of the research pen with the studied cubicles marked in blue for studies 1 & 2 and in red for study 3. MU = milking robot, C = concentrate feeding, circle = mechanical brush, square = trough. Rectangles are cubicles, bottom line is feeder bins.

Data retrieval, processing and storage

There seems to be a favourable generalisation of the potential of computer vision in the research community, as a *plug-and-play* technology; namely that it can be easily deployed and start generating interpretable output. The experiences working with computer vision have taught us that it is not a general truth. Several layers of hardware and software are needed.

The physical setup can be, in the case of this research, split into two parts. The first, is the barn part, with cameras, cabling and switches. The second is the "computer room" hosting a network video recorder (NVR), the computers processing the images, a proxy server for remote access, and a storage unit. The setup went through several iterations, both in the "computer room" and in the barn. During phase I, two switches (Enterprise 8 POE, Ubiquiti, USA) with power over Ethernet (PoE) ports (necessary to power the specific cameras used) were placed on the walls on either side of the pen. RG45 cables were drawn from the switches to the cameras. This meant that a substantial amount of work was necessary to place or move a single camera, as the entire cable (up to 50m) had to be rerouted. The switches were connected to a NVR (Dream Machine, Ubiquiti, USA) in the computer room. The NVR also acted as a cloud gateway to make the live video footage available remotely via the maker's own app. It also routed the frames to the computer to run the pose estimation model. The videos were stored on the hard drive of the NVR. The results from the 3D pose estimation were stored on internal hard drives of the processing computers.

During phase II, a more scalable system had been put in place. Switches directly over the cow pen were connected to the NVR. In addition, ethernet plugs with cables running to the switches had been placed in several locations above the pen, allowing us to easily change camera positions, using cables from 1 to 3m and rarely 5m or 10. The 3D poses and the videos were stored on a NAS storage unit (DS1825+, Synology, Taiwan).

4.2.2 Calibration of the system

The calibration of the system involves three steps:

- Intrinsic calibration to determine each camera's parameters (distortion, focal point).
- Extrinsic calibration of the multi-camera setup to determine intersecting lines of sight and relative location of the cameras to each other.
- Alignment of the multi-camera coordinate system with world coordinates and known origin and axes.

Intrinsic parameters

Intrinsic calibration determines cameras' intrinsic parameters using structure from motion (SFM). These parameters are namely focal lengths, principal point, and lens distortion coefficients, collectively representing the intrinsic matrix K_i. Techniques for determining intrinsics from SFM are not the purpose of this work and only presented here for context.

Cameras record while being moved handheld at a slow pace, along a path describing infinity signs, maintaining the orientation towards a set scene to maintain overlap across frames. The scene must be between 1 and 5m away and contain straight ridges. The process begins with establishing epipolar geometry by detecting ridges and points and matching them across successive frames. This populates the fundamental matrices of feature correspondences between frames. These pairwise relations provided the basis for self-calibration, where intrinsic parameters are refined by enforcing geometric constraints across multiple views of the same static scene. The intrinsic parameters were refined iteratively to achieve consistency across all image pairs.

Extrinsic parameters

For calibration, a scene is recorded from all cameras with a human walking throughout the entirety of the volume to be calibrated. Initially, keypoints on the human (head and joints) are detected on the frames from a pair of cameras. These keypoints establish correspondence between the views of both cameras. Since there is only one of each keypoint (chiral keypoints like shoulders or elbows are labelled as left and right and thus unique), a single correspondence is made for each point on every camera. The system represents a camera by the pinhole model as a single point in a 3D space (in world coordinates). Each camera i is represented by a projection matrix:

$$P_i = K_i[R_i T_i]$$

Where K_i is the intrinsic matrix, R_i the rotation, and T_i the translation. The extrinsic calibration problem is to estimate R_i and T_i. which align the local camera coordinate system with the shared world coordinate system.

In the first phase, each keypoint defines a ray (line of sight) in 3D space that extends from the camera centre through the detected 2D location on the image plane. The system then uses the 8-point algorithm (Hartley, 1997) to determine extrinsic parameters (R_i and T_i) and applies RANSAC across frames to attenuate the influence of noise in the original prediction. Once a pair is calibrated, other cameras are added iteratively in pairs. A preliminary 3D pose is estimated based on the initial parameters.

In the second phase, iterative bundle adjustment optimizes extrinsic parameters and 3D pose. Reprojection error is calculated by retracing each 3D keypoint back onto the cameras and comparing the reprojected point to the original keypoint predicted by 2D pose estimation. Extrinsic parameters are refined by optimizing an objective function that minimizes reprojection error. In addition, it integrates constraints and priors to ensure biomechanical plausibility. This includes notably constraints on joint angles and limb length and penalizes abrupt accelerations or other higher-order derivatives in joint trajectories (Moliner et al., 2021). Optimization uses Huber loss function, with varying weight assigned to the different keypoints' reprojection error depending on prediction confidence and distance (factors empirically linked to estimation accuracy) (Huang & Moliner, 2022).

Aligning to known coordinates

Once the multi-camera system has been calibrated, its coordinate reference can be anchored to a known coordinate system using a calibration plate of known dimensions shown on Figure 6. The plate contains three markers arranged to form a 90° angle, with the distance between each marker fixed at exactly 0.8 m. Each camera detects the markers in its image plane, and then the system triangulates their 3D positions using the previously estimated extrinsic parameters. The coordinates of three reconstructed points are then compared to their known geometric configuration, which serves as a reference frame with orthogonal axes and fixed scale. By applying a rigid transformation, the estimated 3D marker positions are aligned to the known positions of the calibration plate. This allows the entire reconstructed scene, including all camera extrinsic parameters and subsequent 3D poses, to be expressed in absolute world coordinates with known origin, orientation and scale.

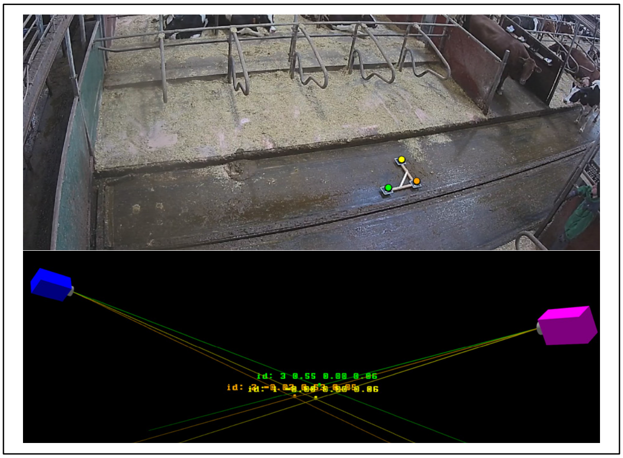


Figure 6. Calibration plate with cameras' lines of sight

4.2.3 3D pose estimation

3D fusion of pose estimation works in three steps: object detection, 2D pose estimation and 3D fusion. The synchronized 2D frames for each scene are processed independently. Rectangular regions of interest are defined on each camera's view, corresponding as best as possible to the areas highlighted on Figure 5. This reduces the processing load by ignoring objects outside the area of interest. Bounding boxes for cows are detected on the frames with YOLOx (Ge et al., 2021) and the contents are passed to the next step. A retrained HRNET identified the location of anatomical landmarks within the contents of these bounding boxes. The nature of convolutional neural networks usually used in image analysis downsamples original images, causing loss of fine-grained detail. HRNet maintains high-resolution representations throughout its entire network, operating several parallel subnetworks at different resolutions and continually exchanging information between them (Wang et al., 2019). A total 24 different keypoints are detected, but the following were used in this work: head at the poll, withers, T13 in the middle of the back, sacrum, at the highest point between the ilia, carpi, and tarsi. This step yields a set of (x, y) coordinates for each frame. Finally, the system established correspondences between the keypoints on each frame. This task involves correctly matching the keypoints to their respective object; that is, identifying which keypoints in one view corresponds to the same unique instance of this anatomical structure in another view when there are several keypoints of the same type in the scene (for example several heads, one head per cow, each head having to be matched to the correct set of limbs and other keypoints). The system combines the known intersecting lines of sight with anatomical constraints (for example, the head is beyond the neck compared to the withers, and the rear limbs are directly under the pelvis) to match the keypoints to the correct object. Once the 2D keypoint correspondences are established, the 3D location of each anatomical keypoint is reconstructed using triangulation. Triangulation involves finding the point in 3D space that, when reprojected against each camera's line of sight, produces a correspondence to the observed 2D locations in each view. The system also performs a temporal filtering step to smooth the 3D pose trajectories over time. This helps to reduce jitter and improve the stability of the pose data. The output of this process is a time series of 3D poses for each object, where each pose consists of a set of 3D keypoints corresponding to the location of anatomical structures.

4.2.4 Data management

The multi-camera system produces 3D coordinates for all 24 keypoints of each object detected in the scene from synchronized video sources. The video is either post-processed, or processed in near-real time, where frames are sent to a processing buffer. This process involves several machines at different locations, and conceptual phases to index, store and analyse the keypoint data.

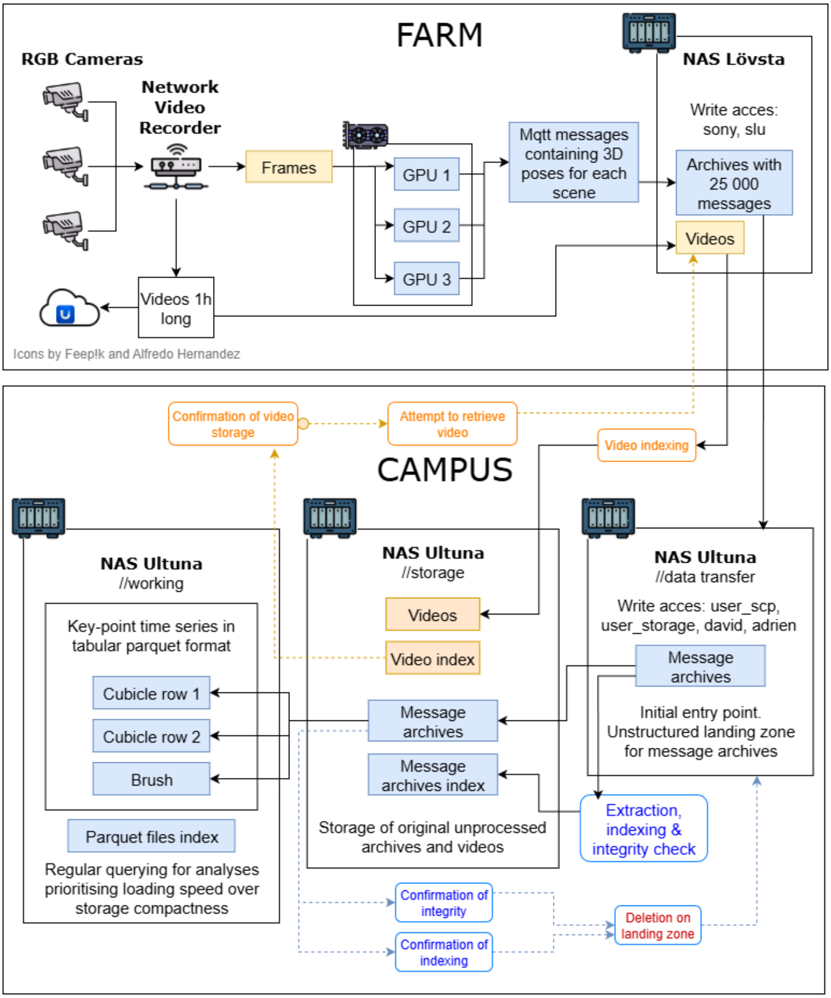


Figure 7. Conceptual architecture of the data pipeline from video acquisition to indexing and long-term storage.

Post-processing synchronized video

In the case of post-processing, pose is estimated on all recorded frames and 3D poses are available at the same framerate as the cameras (29.9 FPS). The output is organised in messages for time windows matching the videos (1 full hour).

Real-time and near-real time pose estimation

In this case, frames are sent from the network video recorder (NVR) for processing. The camera's framerate is 30 FPS. However, in crowded scenes, processing all frames requires more than 0.033s, generating a buffered queue of frames. Computational requirements are also dependant on the number of cameras employed. In order to align sampling rate with processing capacity, the last scene in the processing buffer was processed with pose estimation. This led to occasional variations in sampling rates since crowded frames took longer to process. To maintain synchronization, the first frame with a common timestamp was considered frame 0 for each camera (see section on calibration). All cameras record at the same framerate and the frame's order of arrival in the processing buffer is recorded. All frames from every camera with the same order of arrival are processed together.

The poses for each scene are written into an MQTT message and appended to a JSON file. When the JSON file reaches 25 000 messages, it is compressed and stored. Cameras are re-synchronized before starting to write a new file.

Storage and indexing

A routine was developed to download the archived messages to another storage unit, decompress and index them. The workflow is shown on the lower pane of Figure 1Figure 7

Archives are copied from the remote with rsync onto a landing zone. Each archive's name is compared against entries in an index, identifying new files that were not yet indexed. For each new archive, the script extracts and parses JSON messages contained within. Each message includes a timestamp and one or more detections of objects in 3D. Messages are expanded so that each detection becomes an individual record. These records are then organised into tabular form as Parquet files. The parquet files are stored in a separate query zone. Processed archives are moved from the landing zone to the storage zone.

Downloading archives onto the landing zone is handled as a different step by a user account with writing rights only on that zone. Another user account with reading rights on the landing zone and writing rights on the query zone updates the index, in such a way that no process has at any time both read and write access on a folder that they are accessing, preventing data losses in cases of mistakes or crashes.

The indexing process also records the temporal coverage of each archive by identifying the earliest and latest timestamps within its contents. These timestamps, along with metadata such as the data source (camera group), and file path, are appended to the index. A separate function used for the analyses identifies parquet files of interest based on selected timestamps and according to the time ranges in the index and reads them into Pandas data frames.

Archives that do not contain any valid 3D data are flagged, and a log is generated for traceability. The routine includes versioned backups of the index before modification and differentiate between development and production runs by having a mirrored development index and destination folder.

4.2.5 Data preparation

Defining continuous tracks

After 3D fusion of poses, each detection is assigned a unique id. If a new object is close in space to an object on the previous scene (thresholds for distance unknown), the new object is associated the same id as that of the object on the previous scene. This relates detections pertaining to the same individuals with an unquantified level of confidence and represents a form of pseudo-tracking. Occasionally, when individuals are close to one another, track ids are mistakenly swapped. During real-time processing, the tracking algorithm can only compare the distance between the keypoint in the current scenes and the previous ones. In post-processing, the spatial position of keypoints for a track can be compared to chronologically successive and anterior scenes, and operate on smoothened trajectories increasing robustness. Post-processing also allows to run more demanding tracking algorithms without reducing the framerate.

To identify and rectify these instances, I developed a function that detects sustained spatial discontinuities that operates for cows in cubicles. For each scene, the Euclidian distance between consecutive positions of the withers and sacrum keypoints is calculated. Sudden discontinuities above 0.63cm (half the width of a stall) between scenes (approximately 0.05s) are flagged. The mean position of the keypoints in windows of 5 frames (0.25s) before and after the flagged discontinuity is calculated. If the difference in position persists above the threshold (difference above 0.45cm between the windows before and after the discontinuity), the event is flagged as a track swap. If there is a disruption in spatial continuity that returns to its expected value after 2 scenes and remaining at its expected level, the event is flagged as a noise peak instead. Object ids are then split into several different tracks, with a new unique identifier starting at the index of each id swap. These new ids constitute a form of track. Tracks likely pertaining to the same individual are then merged.

4.2.5.1 Merging tracks pertaining to the same individual

After identifying and segmenting potential id swaps and creating short tracks, objects that pertain to the same individual were merged. First, for each track segment, the mean three-dimensional location of the withers was computed while the animal was positioned within a stall. The stall area boundaries were defined along the horizontal axis using empirically determined limits specific to each camera group. For each track, the mean Y (parallel to the row of beds) and Z coordinates of the withers were calculated. The former coordinate represents the placement in a bed along the row of beds and the later whether the animal was standing or lying.

All tracks were then compared pairwise in terms of their mean Y and Z coordinates and their temporal extent. Pairs of tracks were considered candidates for merging if they (1) occupied spatially close positions in the stall (within 0.63 m in Y and 0.4 m in Z) and (2) either overlapped in time or were separated by a short temporal gap (less than 90 s). These pairwise relationships were represented as edges in an undirected graph, where each node corresponded to a track and each edge connected tracks fulfilling the proximity and temporal criteria. Connected components in this graph were treated as clusters representing a single continuous track. To ensure that no cluster spanned spatially distant beds, an iterative refinement step removed edges whose cumulative spatial separation along the Y-axis exceeded the defined distance threshold. After this refinement, all tracks within a connected component were assigned a common merged track number. To restore temporal continuity between merged tracks, small gaps between merged tracks were interpolated.

Interpolating missing poses

Several events could increase the expected time of 0.033s between frames: crowded scenes may lead to increased latency, failure of the matching algorithm to produce a 3D pose or merging tracks. To ensure temporal consistency, 3D keypoint time series were systematically resampled to 30Hz. Resampling was done using cubic interpolation. In previous work aimed at interpolating cow positions, Ren et al. (2022) had found Akima to be the most faithful method for 2D locations. However, I observed it to regularly overshoot in the vertical dimension compared to expected trajectories and used cubic interpolation instead (which the aforementioned authors had also found satisfactory).

First, consecutive timestamps were examined for irregular temporal gaps. When the interval between two frames exceeded the nominal recording interval (0.033 s) l rows were inserted at regular intervals between them to re-establish a uniform temporal sampling. These inserted rows were flagged as interpolated observations and retained the metadata of their nearest preceding frame (track id, sequence, and camera group). In the second step, the missing coordinates were estimated by spatial interpolation. For each keypoint (e.g., head, withers, sacrum), the available X, Y, and Z coordinates from neighbouring frames were treated as known samples along a one-dimensional temporal axis defined by frame indices. The scipy.griddata fits a local cubic function between adjacent valid observations to estimate smooth intermediate coordinates, and predict the value at the missing observation.

Smoothing and filtering

The 3D fusion process can introduce noise from various sources: vibration of the cameras creating an offset between the calibrated lines of sight and the view, key-point jittering around the ground truth, and erroneous detection. The frequency of the occurrence of each was not quantified but in the methods section of Paper II, you can see how we analysed their effect on the accuracy of an event detector.

Existing literature recommends applying a low-pass filter with a cut-off frequency of 10Hz for behaviour classification tasks (Hamäläinen et al., 2011; Riaboff et al., 2022). These were based on developments with accelerometers but remain sensible in this case. If we assume keypoint jitter to be random and normally distributed around the true position, we can consider that it moves at one frame and then goes back, giving a displacement on 3 frames. At the sampling rate of 30 fps, this gives us a cutoff of 10Hz. Based on this reasoning, I applied a low-pass filter of order 2 with a cut-off frequency of 10Hz in phase I and halved it to 5Hz in Phase II. Nyquist factor was 0.5. I had quantified keypoint jittering by the median 3D Euclidian displacement across frames during stationary phases (before rising). Median jittering was lowest at the head (0.01m, interquartile range (IQR) of 0.009m) and highest at the tarsi (0.07m, IQR = 0.08m). Upon visually inspecting the displacement of the keypoint before and after smoothing, this method preserved the key-point displacement information while filtering out small variations. To attenuate remaining noise, I applied Savintsky-Golay smoothing to the time series of each coordinate of each key-point separately. This technique fits a polynomial in a window centred on each point successively and returns the predicted value of the polynomial at that location. The parameters were 3rd order polynomial to a window length of 15 (0.5s).

4.3 Event detection

4.3.1 Detection of posture transition events

Two different approaches have been proposed to detect posture transition events from sensor output. The first case uses the displacement of anatomical landmarks, used on cows fitted with motion capture reflectors (Ceballos et al., 2004). The rationale behind it, is that the output data "the 3D positions of anatomical structures" are immediately interpretable, in the sense that, for example, a downwards movement of the withers by 0.1cm/s corresponds to a cow transitioning from a standing to a recumbent position. The second method was adapted to accelerometers attached to the leg. When cows are standing, gravity is aligned with the Z axis of the accelerometer, whereas when the cow is lying down, the leg is rotated, and gravity loads mostly onto the x axis. When the main axis load is shifted between these axes for a sustained period (>30s), a posture transition event is flagged (Brouwers et al., 2023b).

The initial method using the vertical displacement of the withers was successful in detecting posture transitions but also flagged mounting behaviours (since the vertical position of the withers does also change through these). A refinement was thus to set a vertical displacement threshold of 40% of the initial position sustained over a period of 30s, or until data points are available, whichever the smallest.

During Phase I, the sequences were already trimmed around posture transition events. Several tracks were present in the snippets, and the detection was applied to identify which track was associated with the cow getting up on lying down. During Phase II, we had continuous data for 3 blocks of 2 successive weeks, each yielding tracks of varying lengths. The need this time was to detect when a posture transition occurred. Based on Phase II, I knew that the detector had a satisfactory sensitivity (88.5%). In the smoothened Z (vertical) series of the withers, events where the keypoint crossed the plane at Z = 1m were flagged as potential posture transitions. Then, the median vertical position in the 30s windows before and after the crossing where compared. Using the same threshold as in Phase I, a sustained 40 % difference was considered a posture transition. A ±30s window with all keypoint coordinates for that track was extracted for later analyses.

4.3.2 Detection of stages of the posture transition

Earlier works on cow posture transitions identified 7 stages (Lidfors, 1989; Schnitzer, 1971). In order to measure the selected indicators of posture transition quality (measuring either duration or displacement), the timing of 5 stages in rising and 4 stages in lying down needed to be known. Paper I serves as a proof of concept, with the aim of showing that we can detect at the first of these stages from the 3D keypoint time series. Each phase has a distinct kinematic pattern, that can be seen on Figure 2 on 43. The stages are the following:

Rising

- Rising on breastbone: when the cow initiates the movement of tucking its front limbs under its body.
- Lunge start position: when the cow has gathered its limbs under its body and performed possible readjustment movements
- Head lunge: the point of furthest extent of the head along the body axis when the cow lunges its head forward to offload weight off the hind limbs.
- Lifting of the rear: with the head extended, the cow steps repeatedly with the hind legs to lift its rear.

- Lifting of the front: the cow then extends its front limbs in succession.
- Standing: the first moment the cow is standing with all 4 legs extended, before it stretches its back.

Lying down

- Initial leg bend: the cow bends one of its front limbs initiating the downwards movement.
- Thoracic limbs touchdown: both carpal joints are in contact with the ground.
- Sacrum descent: the cow starts to lower is rear, marked by an increase in the vertical velocity of the sacrum.
- Recumbent position: the cow is fully lying down.

To detect the stages, the I relied on the interpretability of pose estimation data: that position obtained directly relate to observable phases without needing to integrate or differentiate. The method chosen was change-point detections, which identifies changes in times series without pre-supposing constant properties throughout the series. This method had already proved useful in identifying changes in motion on human subjects (Bastian et al., 2024). A notable advantage is that it does not necessitate a training dataset to relate a signal to a ground truth, thereby reducing the need for annotations. A ground truth is nevertheless needed for validation. Change-point detection was implemented in the python package Ruptures (Truong et al., 2020). The Pelt method requires a penalty parameter and detects several change-points. The annotated posture transition stages were used to refine the penalty and identify the kinematic pattern of the keypoints around the change-point corresponding to that event. A more in-depth description of the method and the parameter search can be found in the *Methods* sections of Paper I and II.

Table 2. Posture transition phases and methods for detection (modified from Paper II).

Posture transition phase	Penalty	Variables for change-point detection	Criteria for selecting a change-point
Rising (LTS)			
Start of rising motion	10	Withers Y, Withers Z	First change point where the median Z withers in the following 1s window > median Z withers in the initial 1s of the sequence
Lunge start	Last point before lunge where withers forward velocity = 0		
Head lunge	Return to head velocity = 0 after highest peak		
Standing	5	Withers velocity	First change point after the last velocity peak of 0.18 (normalized units)
Lying down (STL)			
Initial leg bend	10	Withers vertical velocity	Last change point before the first peak in withers velocity above 0.2 (normalized units)
Thoracic limbs touchdown	3	Withers Z	First change-point immediately after the first peak above 0.2
Sacrum descent	Random forest		
Recumbent position	10	Withers Y, Withers Z	Last change point where the median Z withers in the following 1s window < median Z withers in the final 1s of the sequence

4.3.3 Creation of a ground truth and validation

For Phase 1, the ground truth was annotations on the timing of the selected stages from video. For Paper I, 3 observers annotated the timestamp of the first stage of the rising motion in 60 rising events randomly sampled from the 471 complete sequences (sequences without interruptions in the objects ids). Thirty sequences were common to all observers, 10 were unique to each observer and 15 were randomly resampled within each observers own set to measure intra-observer consistency. The sequences were blinded and shuffled. Observers were given the videos for these sequences from all 6 or 7 camera angles and provided with the following definition: *The cow is lying down and rises on its breastbone and elbows, which causes the withers to* *rise visibly above the rest of the back.* Observers then calibrated together by agreeing on the time to annotate from 5 training sequences different from the validation set.

For the remainder of the events, labelled for paper II, each observer annotated 100 sequences for both rising and lying down events, of which 55 (per posture transition type) were common to all observers, 30 were unique, and 15 were randomly resampled. 10 different sequences were used for calibration of the observers. The exact definition of the phases provided to the observers can be found in Table 1 of Paper II.

For Paper III, observers labelled the individual cow performing each bout in the selected sequences. Identification sheets were developed, with images of each individual cow from various angles. For each sequence, one of either two observers annotated which cow was performing the bout, and the cubicle number. The cubicles were counted 1 to 6 for each row separately, starting from the left when facing the front end. If an observer was unsure (for instance because of an even coat with few distinguishing patterns), they flagged the annotation as such. They left the annotation blank if they could not identify the individual. These sequences were excluded from the analysis.

4.4 Scoring of posture transition indicators

Both scientific literature, and industry guidelines recognize sets of measures that are used to assess the quality of cubicles when it comes to allowing for comfortable posture transitions. A measure commonly found is the “lunge space”, and by extension the “bob-room” (Cook, 2019). This is a measure of the space available in front of the cubicle, both forward and upward for the cow to extend its head forward during the lunge movement. We have seen in Section 0 that the head lunge is an innate and biomechanically important movement. It serves as a way of displacing the cow’s weight forward when getting up, with the effect of relieving some of the body’s load off of the rear limbs, reducing the effort needed to lift the rear (Lidfors, 1989) (see Figure 3). The lack of lunge space is an issue raised with wall-facing cubicle models, and the resulting lunge movements to the side are characterized as an abnormal motion (Brouwers et al., 2024). More recent barn designs allow for space in front of the cubicle, however, other elements, such as a head rail, represent a forward barrier and lead to collisions during rising motions (Veissier et al., 2004). As a result, estimating usable lunge space is not as trivial as measuring the space in front of the cubicle. This brings us to the usefulness of pose estimation in 3D.

Table 3. Selected indicators of posture transition quality

Rising	
Duration of rising motion	To avoid collisions with the metal bars, cows take slower more hesitant movements, which result in longer bouts.
Backwards crawling on carpal joints	The cows lie down under the neck and head rail. The fast-rising motion risks collisions with the bars (Veissier et al., 2004). The cows therefore crawl backwards to increase the upwards and forward space before rising. Contact with the straps may not be perceived as adversely as with the bars, reducing the need for backwards crawling.
Delayed rising	Readjustment motions that delay the rising motion are a way for the cow to cope with a restrictive environment by positioning its body before rising. Flexible straps with less adverse contact may shorten this phase.
Head lunge distance	The rigid head rail represents a physical limit to forward lunge, whereas the flexible strap can, to some extent, move forward with the cow's body.
Head "bob"	During lunge, the cow should be able to "bob" its head up and down (Cook, 2009), flexible straps allow for greater amplitude by acting as a soft boundary.
Side lunge	Side lunge is seen as a compensatory mechanism when forward lunge space is perceived to be insufficient (Cook & Nordlund, 2005). Cows can push against the strap when lunging, reducing the need to lunge to the side.
Lying down	
Duration of lying-down motion	Comparably to rising, the cow lies down slowly to avoid hitting rigid structures. Flexible straps are expected to increase movement swiftness.
Hind quarter shifting	Lowered risk of collision decreases hesitation and readjustment movements.
Head displacement	Flexibility of the head rail reduced aversion to contact and favours further extension.

Animal-based measures are preferred for welfare assessment and constitute the majority of indicators in WQ (Blokhuis et al., 2013). The assessment in WQ is visual, and accurately quantifying lunge distance is unfeasible. A visual indicator to estimate whether lunge space is perceived as insufficient is side lunge (Brouwers et al., 2024). However, a study aimed at automating the classification of side versus straight lunge ran into many ambiguous "edge cases" of "slightly sideways lunge", showcasing that a dichotomous indicator, practical for visual observations, might not reflect a continuous reality (Brouwers et al., 2023b).

Using this knowledge, I attempted to compile existing measures and define them quantitatively rather than classifying abnormal versus normal. These measures are selected based on the following criteria: (i) animal-based measures (ii) used in previous studies evaluating cubicle comfort through cow movements, (iii) can be measured using pose estimation in 3D, and (iv) are expected to be affected by cubicle design. We will refer to these measures that relate to the quality of the posture transition as indicators of posture transition comfort. The indicators are listed in Table 3. The way they were calculated differed slightly between Phases I and II of data collection and the exact method can be found in the accompanying articles (Papers II and III) at the end of this thesis.

4.5 Intervention study

The intervention sought to compare movement patterns and indicators of comfortable posture transition between cubicles with rigid (metal) head and neck rails versus cubicles with flexible straps. A quasi-experiment was conducted where 12 of the cubicles were changes to the flexible design during two weeks and indicators compared between cubicle designs.

4.5.1 Hypothesis development

The aim of the experiment is to test whether cows are more comfortable transitioning between posture in cubicles with flexible head and neck rails that in cubicles with rigid metal bars. To test this, we would ideally posit the following hypothetico-deductive (HD) model, which is exemplified with lunge distance but works in the same way for the other indicators and their interactions:

- Hypothesis (H): Cows are more comfortable getting up in cubicles with flexible head and neck straps.
- Auxiliary assumption (A): when cows are comfortable, they lunge further and straighter.
- Expected observation (O): cows in flexible cubicles use more lunge distance than in cubicles with rigid bars.

The issue with the model above is the limited available evidence for the auxiliary assumption, although it is highly plausible. The existing evidence-base regarding lunge distance states that cows in more permissive environments (larger cubicles, open packs or pasture) lunge further and exhibit generally more fluid rising motions (Brouwers et al., 2024) but does not guarantee the immediate corollary: that cows showing more fluid and ample motions are in a less restrictive environment and automatically more comfortable. Most of the evidence for a causal relationship between cubicle design and lunge room concerns restricted environments leading to insufficient lunge room, and the link with comfort is not trivial. Regarding the other indicators, it also states that cows take longer time getting up, with more hesitation and abnormal movements.

Thus, it is more appropriate, given the available evidence for the link between comfort and restrictive cubicles, to frame the HD model in terms of discomfort, then interpreting a reduction in discomfort-related behaviours as an increase in comfort:

- Hypothesis (H): Cows rising in cubicles with flexible head and neck straps experience less discomfort compared to those in cubicles with rigid metal bars (all other cubicle features equal).
- Auxiliary assumption (A1): Rigid metal bars limit spatial use and forward head movements.
- Auxiliary assumption (A2): Limiting movement opportunities interferes with the cow's natural rising kinematics, causing collisions, hindering ability to balance weight, and altered rising patterns, notably shorter or sideways head lunge.
- Auxiliary assumption (A3): Collisions and disruption of normal rising behaviour create discomfort, which may manifest as increased effort, hesitation, aversion, or stress.
- Corollary assumption (C): the absence of restriction allows for more freedom of movement limiting the risk of the aforementioned externalities.
- Expected observation (O): cows in flexible cubicles use more forward space than in cubicles with rigid bars.

A1 and A2 are supported by biomechanical studies of cows rising (Brouwers et al., 2024; Ceballos et al., 2004) and state fairly logically that physical barriers in the way of cows' spontaneous, motion patterns will inevitably induce a change in the movement. When these barriers are in the way of normal movements, collisions happen (Zambelis et al., 2019). A3 builds on the premise that disruptions in innate motion patterns are associated with negative affect (Lidfors, 1989), proposing that when rising becomes more difficult, or abnormal due to restriction, cows experience increased discomfort (Nielsen et al., 2023). Corollary assumption C follows from the preceding logic: that reducing physical barriers allows for more freedom of movement and thus a reduction in the adverse experiences associated with constrained movements. Together, the assumptions state that restrictive cubicles form a physical barrier to rising motions, that this barrier leads to collisions and reduced head lunge, and finally that the adverse experiences create discomfort. We deduce conversely that increased head lunge is a visible sign of a less restrictive environment and thus less discomfort.

4.5.2 Experimental design

The experiment follows an intervention quasi-experimental observation study design, with each animal serving as its own control. Pose estimation in 3D, and location (from the ear tag) were collected in 12 of the cubicles (red area in Figure 5) out of 65 total in the pen. Cows had access to all cubicles throughout the experiment but only their bouts occurring in the 12 cubicles were recorded. This means in practice that cows could chose to lie down in control cubicles even during the intervention stage.

The control cubicles and their dimensions can be seen on the upper panes of Figure 4. They consist of a bed, dividers and a head and neck rail. The experimental cubicles are the same with the head and neck rails replaced with flexible straps (CC1800 with flexible front and neck bands, DeLaval, Sweden). These can be seen on the lower pane of Figure 4.

Based on the results of data collection Phase I, I estimated being able to collect at least 800 true positive sequences of each posture transition with sufficient data quality (uninterrupted tracks and low noise) in a period of 32 days using 12 cubicles. There was a standard deviation in head lunge distance of 0.33 (arbitrary spatial units close to the meter but of unknown uncertainty). A power calculation adjusted for 6 intra-individual repetitions with an intra-individual correlation of 0.46 revealed a necessary sample size of 41 animals to find a statistically significant difference in head lunge distance of 0.087m. This represented 249 events or 10 days. This duration was increased to two weeks. More details on the power calculation can be found in the *Methods* section of Paper III.

3D poses and location were collected continuously for 14 days in the control cubicles. The head and neck rails were replaced with the flexible straps. The holders for the flexible straps had been installed in anticipation to reduce intervention time strictly, which took two workers about a half workday. Cows were given a 7-day adaptation period before data was collected again for 14 days. Then, the rigid bars were re-installed, and after an adaptation period of 7 days, data was recorded for 14 more days.

4.6 Statistical analyses

The data in Phase I was strictly observational; we recorded bouts without intervention on the cows that could modify the bout. In this phase, I sought to test the association between indicators. After the intervention study in Phase II, the marginal effect of the intervention on the indicators was tested in a mixed effects model:

$$y_{i,n} = \beta_0 + \beta_1 * flexible_{i,n} + \beta_2 * Group_i + u_i + \varepsilon_{i,n} \qquad \text{Equation 1}$$

Where y is the indicator value for cow i at observation n, $Group$ is a binary indicating the camera group to account for differences in location of the calibrated origin, and u is a random effect for cow. Type I risk was set at $\alpha = 0.05$ with Benjamini–Hochberg correction to account for the testing of several potentially correlated indicators. This works by ranking p values from smallest to highest and assigning each p-value a threshold $p_i \leq \frac{i}{m} * \alpha$ where i is the rank (from smallest to largest p-value) and m the number of tests. The correction was applied within each bout and analysis type so that $m_{rising} = 7$ and $m_{lying} = 4$. The following indicators were box-cox transformed before testing: lunge angle, rising delay, backwards crawling distance, shifting duration and head displacement when lying down.

The change in indicator distributions between baseline and flexible cubicle configurations was analysed differently, depending on the indicator's distribution found in Phase I.

Backwards crawling and rising delay during LTS and shifting duration during STL had excessive zeros and were analysed in two steps with a hurdle model. For the zero component of the model, a logistic regression tested the effect of the flexible configuration on the probability of the indicator being greater than 0 according to equation 14:

$$logit(P(y_{i,n} > 0) = \beta_0 + \beta_1 * flexible_{i,n} + \beta_2 * Group_i + u_i + \varepsilon_{i,n} \qquad \text{Equation 2}$$

For the continuous part, a subset of the samples was created excluding events with $y_{i,n} = 0$. A mixed effects model was fit according to the equation below, to test the effect of flexible configuration on the strictly positive part of the distribution:

$$y_{i,n} = \beta_0 + \beta_1 * flexible_{i,n_{y>0}} + u_i + \varepsilon_{i,n_{y>0}} \qquad \text{Equation 3}$$

The zero component provides information on the likelihood of observing a null result, while the continuous component provides information on the effect size on non-zero events.

4.7 Force distribution modelling

After a preliminary analysis of the distribution of indicators, without correcting for individual variation at this stage, we were surprised to notice that there were no overall differences in the duration and spatial use, especially at the head. This prompted a further investigation; either there was truly no overall difference, either the experimental design was flawed, or we were not looking in the right direction. The experimental design might have presented a limitation, but available literature on flexible neck rails, and our own observations of the video point towards differences between rigid and flexible head and neck rails. To determine the likelihood of there being no effect, we had to rule out the fact that we had looked at the wrong indicators.

I went back to the theory behind rising motions. The lunge motion acts as a way of balancing the weight between the front and the rear (Lidfors, 1989). Observations of rising bouts in too small cubicles noted unsuccessful rising attempts (Tschanz & Kämmer, 1979), or attempts where abnormal strain placed on the limbs lead to skin and muscle lesions (Kohli, 1987). This produced the subsequent research question: “can we model the shift in weight distribution throughout the bout?”. If possible, we would effectively quantify the biomechanical driver behind the lunge motion, rather than an observable consequence.

Insights from the field of biomechanics showed that modelling forces imposed on joints, and ground reaction forces using motion capture was an accepted method, at least in human biomechanics (for example Johnson et al., 2018). Translating this method to pose estimation in 3D could allow to model forces during rising motions.

I proposed a simplified model, representing the cow as a set of two rods between the sacrum and withers and the withers to the head. These rods are supported at the sacrum by the rear limbs and at the front by the forelimbs. The analysis was restricted to the lunge phase of rising between T_{lunge_start} and T_{lunge_max}, during which the front limbs are folded under the body and the carpi act as the front support (see Figure 8). During this stage, the cow extends its rear limbs to lifts its rear. The centre of mass (COM) was set at 55% along the sacrum to withers rod, based on previous work (with load cells) finding that on average, 55% of the body weight is supported by the front limbs (Chapinal et al., 2009; Neveux et al., 2006). The model was two-dimensional; at each time-point, a plane was defined collinear to the horizontal and to the withers-sacrum axis. The plane is defined by the orthogonal vectors $(0,0,1)$ and $(x_{withers} - x_{sacrum}, y_{withers} - y_{sacrum}, 0)$ at the timepoint *start of lunge motion*. The rationale was that we rotate the y axis so that it becomes longitudinal to the cow's body. The midpoint of the carpi projected onto this plane forms the front support point coordinates: $(\frac{y'_{lcarpus}+y'_{rcarpus}}{2}, \frac{z_{rcarpus}+z_{lcarpus}}{2})$ where *x*' and *y*' are the reprojected coordinates and subscripts r and l refer to right and left. Because of occlusion in the claw keypoints keypoint, the rear-support was estimated by projecting the tarsi onto the ground so that the rear support is defined by the coordinates: $(\frac{y'_{rtarsus}+y'_{ltarsus}}{2}, \frac{z_{rcarpus}+z_{lcarpus}}{2})$. Note that the vertical coordinate of the carpus is intentionally used here because they are in contact with the ground at this stage and thus represent the best estimate of the ground location.

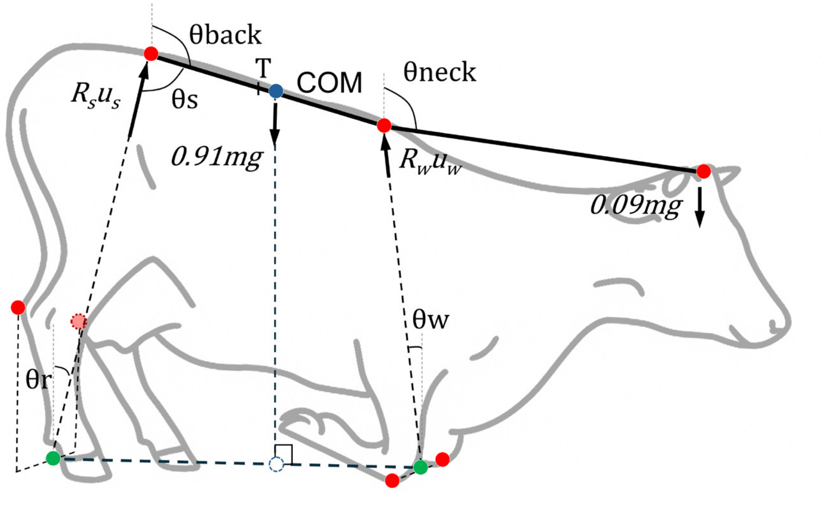


Figure 8. Simplified force loading model

In the first stage of the analysis, I modelled the forward displacement of the COM projected onto the support-to-support axis (dashed blue vertical line on Figure 8). The rationale is that when the rear limbs are placed closer to the front limbs, they support more weight and lie closer to the centre of gravity. Because the cubicle limits the available longitudinal space, especially for larger cows, the rear limbs may not be able to extend as far forward as they would otherwise. Therefore, projecting the COM onto the support-to-support axis (rather than using its absolute forward displacement) provides an estimate of how far the cow can shift its centre of gravity away from the hind limbs, given their actual position. The maximum COM shift during the lunge phase was extracted and analysed with the mixed-effects model in Equation 1. To investigate whether the effect of flexible straps was proportionate to baseline displacement (i.e., whether cows with lower baseline displacement showed a greater predicted increase than those already showing a large displacement), the following model was used:

$$y_{i,intervention} - \mu_{baseline_i} = \beta_0 + \beta_3 * \mu_{baseline_i} + \beta_2 * Group_i + u_i + \varepsilon_{i,n} \quad \text{Equation 4}$$

Where μ is the mean COM forward shift at baseline for individual i.

This first step provides an initial static estimate as to how weight can be displaced. Translational and rotational acceleration will also affect the forces on the supports. To that end, we (as in myself together with an expert in biomechanics) developed the simplified force loading model presented in

Figure 8. Forward velocity and acceleration of the withers are calculated by the change in successive observed positions of the smoothened keypoints as $v_W = \frac{W_{x',t} - W_{x',t-1}}{\Delta t}$ and $a_W = \frac{(v_{W,t} - v_{W,t-1})}{\Delta t}$ where x' is the x coordinate reprojected against the back axis, *t* is the timestamp of one specific frame and *Δt* is the time difference between two successive frames equal to ${}^{1}/_{30}$s. The dynamic equilibrium of the forces is posited in its horizontal and vertical dimensions respectively as:

$$R_S \cos\theta_r + R_W \cos\theta_W = 0.91 m a_{T,x'} + 0.09 m a_{H,x'} \qquad \text{Equation 5}$$

$$R_S \sin\theta_r + R_W \sin\theta_W = 0.91 m a_{T,z} + 0.09 m a_{H,z} + (0.91 + 0.09) g$$

Equation 6

where *g* is gravity, subscript *H* is the head keypoint and *T* is the midpoint of the back. Moments were calculated about the withers. The sacrum and head generate moment through their lever arms, the contribution of the head to the front reaction is expressed through translational acceleration and torque exerted at the withers pivot point. The moment about the withers is:

$$-R_S L_{back} \sin\theta_S \qquad \textit{[reaction moments]}$$

$$+0.91 m g L_{W\to T} \sin\theta_{back} - 0.09 m g L_{H\to W} \sin\theta_{neck} \qquad \textit{[gravity moments } M_g\textit{]}$$

$$-\ 0.91 m L_{W\to T}\left(\sin\theta_{back} a_{T,z} + \cos\theta_{back} a_{T,x'}\right) \qquad \textit{[back rotation } \tau_T\textit{]}$$

$$+\ 0.09 m L_{W\to H}\left(\sin\theta_{neck} a_{H,z} - \cos\theta_{neck} a_{H,x'}\right) \qquad \textit{[head rotation } \tau_H\textit{]}$$

$$= I_{back}\alpha_T + I_H \alpha_H \qquad \textit{[inertia]}$$

Equation 7

where *I* is the inertia of a segment, *L* is its length, *a* is the acceleration and α the angular acceleration. All other subscripts can be found on Figure 8. I_{back} is assumed as $\frac{1}{12} m L^2$ where L is the length of the segment and $I_H \approx 0$ with the mass concentrated in the head. We split the ground reaction forces into their vertical and horizontal components and obtain:

$$\begin{bmatrix} \cos\theta_s & \cos\theta_w \\ \sin\theta_s & \sin\theta_w \\ -L_{back}\sin(\theta_S) & 0 \end{bmatrix} \begin{bmatrix} R_{s,dyn} \\ R_{w,dyn} \end{bmatrix} = \begin{bmatrix} 0.91 m a_{T,x'} + 0.09 m a_{H,x'} \\ 0.91 m a_{T,z} + 0.09 m a_{H,z} + mg \\ I_{back}\alpha_T - M_g - \tau_T - \tau_H \end{bmatrix}$$

Equation 8

This is noted as $\mathbf{A}\begin{bmatrix} R_s \\ R_w \end{bmatrix} = \mathbf{y}$. Since it is over-constrained (1 degree of freedom remaining from 3 equations to solve 2 reaction forces), we solve for

R_S using least squares $\begin{bmatrix} R_s \\ R_w \end{bmatrix} = \left(\mathbf{A}^{\mathrm{T}}\mathbf{A}\right)^{-1}\mathbf{A}^{\mathrm{T}}\mathbf{y}$. The effort placed on the rear limbs was quantified as the cumulative work throughout the lunge motion Total work of the rear limb throughout the bout was defined as the cumulative displacement times the instantaneous rear limb forces: $\int R_{S,dyn}\cos\theta_r * dSacrum_z + \int R_{S,dyn}\sin\theta_r * dSacrum_y$

5. Results and discussion

This section presents the main results that will answer the aims of the thesis. The immediate implication of the results is also discussed in this section. More objective and systematic reporting of the results can be found in the papers.

5.1 Detection of posture transition events from continuous pose estimation data

In Phase I, 979 rising bouts and 1015 lying down bouts were "manually" selected. Synchronized video snippets were extracted at ±15s around the posture transition event and 3D pose was estimated for these sequences. Each sequence resulted in several tracks including that of the cow rising or lying down. Applying the detector based on the change in relative positions of the sacrum and withers, 814 and 798 tracks were classified as rising and lying down respectively, equating to false negative rates of 16.9% and 21.4% respectively. 5 and 26 sequences respectively were wrongly classified. This equates to sensitivities of 83.1% and 78.7%.

In Paper I, only 493 sequences where the posture transition was captured in a single uninterrupted track were analysed. In Paper II, sharing the same dataset, we sought to test the robustness to discontinuous tracks and to missing poses (track interrupted for a few scenes). This provided an additional 305 rising and 301 lying down sequences or 37.7% and 38.1% of the total events analysed.

After detecting the posture transition, the detection of each of the stages shown in Figure 2 was necessary to set time bounds within which the indicators would be measured. In Paper I we only detected the first phase, to offer a proof of concept, and an indication that movements in the keypoints did properly capture kinematically meaningful information. Agreement was measured by intraclass correlation (ICC) of a model predicting event time based on observer. The ICC was 0.85 between human observers and 0.81 when adding an effect for machine detection, which we interpreted as similarly acceptable. Disagreement between observers ranged from 0.9s to 1.7s and between observers and machine from 1.0s to 1.3s.

Appending discontinuous tracks (*stitching*) had a significant effect on the accuracy of the automated detection (compared with human annotations) for the stages *rise on breastbone* of rising (-1.4s difference between human and machine) and *thoracic limbs touchdown* in lying down (-0.5s). Interestingly, the first stage, *rise on breastbone* was the most ambiguous to human observers (1 to 1.8s average difference between observers), which might reflect a variability in kinematic profiles making it difficult to both annotate and find a rule flagging the event. Interpolating poses did not have a significant effect on the agreement between human and machine. This was an important finding with implications for continuous monitoring, since we would rely on this data processing methods in the next phase.

In phase II, the added challenge was to detect posture transition in continuous keypoint data. I did not perform an estimation of the false positive rate as this was not the scope of the study. However, the lower sample size at the return to baseline (190 versus 285) could be an indication that the system got less performant with time. During the experimental period, the 5-95 IQR of acquisition rates was 6.5 to 30Hz and the median 30Hz, meaning that the pose estimator was in the vast majority of cases able to keep up with the arrival of frames in the processing buffer and generated mqtt messages providing a proof of concept for real-time implementation.

Altogether, 850 rising sequences and 853 lying sequences were detected. Out of these, 733 rising sequences contained the entire bout, with true detections at the events of interest (not missing detection at key stages) and 787 lying down sequences. For 733 valid rising sequences, 4 had tracks wrongfully stitched that belonged to different individuals on the same scene. This happened for 6 out of 787 valid lying down sequences. The initial power analysis used to determine the number of samples necessary for Paper III based on the true positive rate and variance found in paper I required a sample size of 249 bouts which we estimated to capture in 10 days (25 bouts/day). We captured up to 289 bouts in 14 days (21 bouts/day) from which we can grossly infer a slightly higher false negative rate in detecting posture transitions from continuous data compared with the curated sequences.

5.2 Measuring comfort in cubicles with automated indicators.

In Paper II, we reported the number of bouts exceeding accepted thresholds for comfortable posture transitions. Thresholds for rising were found in the literature or derived from industry guidelines for rising delay

(10s), backwards crawling (0m) and a resource-based value for lunge space (0.6m to 0.9m). Thresholds for lying down were found for total duration (6.3s), shifting duration (3s) and an empirical measure of head displacement in unrestricted environments was also found (0.59m). Altogether, we found that 59.9% of rising and 29.1% of lying down movements were abnormal. Out of all rising bouts, 2.8% exceeded the threshold for rising delay (30.2% if we apply the more conservative threshold of 5s found in the industry framework *Fråga Kon*). Backwards crawling was higher than 0 in 58.2% of bouts. Regarding lying down, 28.9 % of bouts exceeded the threshold for total duration and 8.3 % for shifting duration.

The PCA in Paper II suggested that posture-transition quality cannot be faithfully summarised into a single dimension. Indicators rising delay and total duration were highly correlated (r=0.88), which is expectable as they are nested durations. Backwards crawling and rising delay were moderately correlated (r=0.46), suggesting that duration of the preparation phase (as done in *Fråga Kon*) is a sound summarisation but not an "iceberg indicator". The PCA showed different uncorrelated strategies, combining for instance short lunge distance with swift movements. Extended crawling does not necessarily predict increased effective lunge distance (contrary to initial hypotheses). The components were interpreted as *duration, straight lunge*, *spatial use,* and *fast crawling*. Together, the indicators represent distinct, combinable rising strategies, some of which are considered atypical, and some combinations of both desirable and atypical motions in the same bout. These correlation patterns suggest that multiple indicators are required to describe posture-transition comfort, as each captures a different biomechanical or behavioural adaptation. Strategies might be driven by cow size, which is associated with externalities (Zambelis et al., 2019). For rising, there were only three indicators, with 2 nested durations being correlated, limiting the conclusions available.

The comparison between metal bars versus flexible straps revealed head lunge angle (+2.7° ± 1.0 and +2.7° ± 1.1) and head bobbing space (+0.10m ± 0.03 and +0.14m ± 0.03) to be significantly higher in the flexible configuration compared to both baseline and return to baseline respectively. This supports previous results stating that metal bars limit forward space use and modified neck trajectories (Veissier et al., 2004). There was a significant decrease in duration upon return to baseline (-0.8s ±0.3). Results regarding spatio-temporal use are compiled in Table 4.

There was also a significant effect of the flexible straps on the forward displacement of the centre of mass along the support-to-support axis (2.8% ± 1.0 to 4.7% ± 1.1, $p < 0.05$) and on the maximum offloading of the rear limbs (1.9% ± 0.9 to 3.1% ± 1.0 $p < 0.05$). Increased forward spatial use together with differences in the offloading of the rear limbs increases the available evidence supporting that cows are able to make use of the increased forward movement opportunities offered by flexible straps when getting up (Brouwers et al., 2025).

When we model the changes from baseline to intervention only and add the mean baseline COM displacement in the model explaining changes in COM displacement, we find that the effect of mean baseline COM shift on the predicted COM shift at the intervention is negative and significant ($p < 0.001$). The scale is 0.29 - (0.80±0.12) * (baseline COM shift). This means that cows with a larger initial shift will have a lower predicted increase under the intervention than cows with a low initial shift. For example, a cow that shifts by 30% under baseline will have a predicted increase under flexible straps by 5% while a cow with a baseline forward shift of 15% will have a predicted increase by 17%.

Regarding lying down, the interpretation was less straightforward. There was a significant decrease in total duration between the baseline and the flexible straps that did not return when putting the metal bars back into place. This could be interpreted as less influence of metal head and neck rails on the lying down movement than on lying down, which is supported by earlier findings suggesting that elements in the front of the cubicle represent more of an impediment when rising than lying down (Schnitzer, 1971). The decrease in forward head displacement observed upon returning to metal bars (-0.07m ± 0.03 $p = 0.003$) is consistent with earlier findings (Ceballos et al., 2004) but it is intriguing that it was not observed at the baseline. My current hypothesis as per the reason is that flexible straps allowed the cows to move in further into the cubicle before lying down. This would not have hindered their movements in the presence of flexible straps. However, having adjusted to this new positioning could have decreased their ability to move forward because of the rigid bars in the way. This will be investigated in an upcoming paper.

Table 4. Predicted marginal differences in indicator value between the intervention (2) as reference level and the two baseline stages (1 and 3), order by adjusted α limit necessary to accept the hypothesis that the difference is greater than null. Significant coefficients at adjusted alpha are bolded. Modified from Paper III.

Indicator	Adjusted α limit	Effect* stage 2→1	p2→1	Effect* stage 2→3	p2→3	Pow **	ICC
Rising							
Head "bob" (m)	0.0083	**-0.10±0.03**	**0.001**	**-0.14±0.03**	**<0.001**	0.98	0.29
Lunge angle (°)	0.016	**-2.7±1.0**	**0.014**	**-2.7±1.1**	**0.010**	0.72	0.11
Duration (s)	0.025	-0.4±0.3	0.148	**-0.8±0.3**	**0.010**	0.77	0.49
Backwards crawling (m)	0.033					0	0.23
(Zeros)		0.19±0.24	0.440	0.01±0.27	0.962		
(Non-zero)		-0.01±0.02	0.578	-0.02±0.02	0.071		
Rising delay	0.042					<0.1	0.47
(Zeros)		-0.4±0.3	0.191	-0.1±0.3	0.844		
(Non-zero)		-0.1±0.2	0.824	-0.5±0.3	0.096		
Lunge distance (m)	0.05	0.00±0.02	0.742	0.010±0.02	0.497	<0.1	0.22
Lying down							
Total duration (s)	0.016	**0.6±1.2**	**0.009**	0.4±0.2	0.073	0.91	0.30
Head displacement (m)	0.033	-0.03±0.02	0.167	**-0.07±0.03**	**0.011**	0.70	0.12
Shifting (s)	0.05					<0.1	0.29
(Zeros)		-0.1±0.3	0.66	-0.5±0.4	0.187		
(Non-zero)		**0.3±0.1**	**0.024**	0.1±0.2	0.455		

*For coefficients on which a box-cox transformation was applied, we report the difference in mean prediction between both levels of experiment stage in lieu of coefficient. Zero part reported log odds.

** Power estimated through Monte-Carlo simulation (with 1000 replications). The values represent the lower bound of the 95% confidence interval of the estimated power.

The results from force modelling were inconclusive. There was no significant effect of flexible straps on the amount of work at the rear limbs in the lunge phase. The model was sensitive to changes on the choice of support points for example (projection onto the ground or claw keypoints) and to rotation of the coordinates along the cow axis. There was also a considerable residual term in the residuals of the least-squares estimate of the reaction vertical forces. The median root mean square error of the least squares estimate of reaction forces averaged across events was $0.41 m.s^{-2}$ and the 95^{th} percentile was $1.0\ m.s^{-2}$. This represents an upper range of the error amounting of approximately 11% of gravity. The means that force modelling can be a possible indicator of movement opportunities in stalls but that the work presented here offers a proof of concept rather than a definitive method.

6. General discussion and roadmap

The original vision for the thesis had been to automate a range of animal welfare assessment indicators, which was perhaps overly ambitious. A review on the potential to automate WQ indicators had after all proposed combinations of different sensors depending on the indicator (Maroto Molina et al., 2020). What this project prompted however, was a discovery on the place of pose estimation in 3D for evaluating indicators of welfare for which the motion of anatomical features informs on ease of movement, beyond simply reproducing existing indicators. In this last section, I want to discuss the findings on the improvement of posture transition comfort with flexible cubicles, on what 3D pose estimation can deliver for welfare assessment, and how the notion of welfare assessment is approached when it is automated.

6.1 Assessment of posture transition as a welfare parameter

There will be an inevitable trade-off in cubicle design, notably between cleanliness and comfort (Gieseke et al., 2020). I proposed earlier in the ethical statement, that the goal of assessing welfare was not to evaluate how bad the trade-off was, but rather to find out in which conditions the animals fared the best. From the methods it seems that pose estimation in 3D was able to extract answers to this question. From the results it seems that they fare better in cubicles with flexible straps.

Welfare can be evaluated as the animal's capacity to cope and adapt (physically and mentally) to its environment (Arndt et al., 2022; Broom, 1996). Linking this interpretation to cubicle design, we can evaluate how well the animal is able to adapt to such a system, despite the restrictions they impose. A cow that can rise and lie down fluidly, with sufficient space and without repeated contacts with hard surfaces, is able to meet a strong behavioural need (lying) without excessive physical and psychological cost. Conversely, a cow that anticipates pain or instability during transitions may develop avoidance strategies, such as lying down less frequently (Haley et al., 2000), displaying staggered motions (Brouwers et al., 2024), side lunge (Brouwers et al., 2023b) or delayed rising (Zambelis et al., 2019) and lying down (Gieseke et al., 2020). These may constitute coping but at the expense of comfort and physical health in the form of skin lesions (Zaffino Heyerhoff et al., 2014). Repeated negative experiences during posture transitions are therefore likely to contribute to negative affective states such asa frustration or anticipatory discomfort, while limiting opportunities for positive experiences associated with comfortable rest (Nielsen et al., 2023).

The latest update to the five domains of animal welfare presents mental states as an aggregate of positive and negative experiences arising from the first four domains (Nutrition, Physical Environment, Health, and Behavioural Interactions). This means that what matters for the animal is its own perception of its condition (Mellor et al., 2020). Yet in practice the results presented in Paper II and III are overall means and marginal effects of flexible elements. Welfare Quality prescribes observations on a third of the animals, which is empirically a good estimate of the state of the herd (Blokhuis et al., 2013). Aggregating welfare indicators such as posture transition comfort at herd level represents a practical utilitarian stance that accepts higher burdens on some individuals (Sandøe et al., 2019). At the same time, externalities with cubicle designs differ across individuals for example the prevalence of injuries is associated with cow size, the direction of the effect suggesting that larger cows have more difficulty coping with cubicles (Zambelis et al., 2019). When annotating video, we observed some specific cows to particularly struggle with their rising motions, with examples of slipping and falling. While aggregates provide practical insights, the question at hand is how group-level improvements in posture transition comfort translate to individual experiences.

Work on both cow limb trajectories (Leclercq et al., 2024; 2025) and training of heifers (Paranhos Da Costa et al., 2021) has shown that inter-individual variability can be high, possibly exceeding treatment effects in some cases (Paranhos Da Costa et al., 2021). We found the highest yet moderate ICC in rising duration (used as an "iceberg indicator" in *Fråga Kon*), reinforcing the importance of individual variability. This highlights the importance of understanding each animal's own range of motions, and what might be considered "normal" for one individual (Tijssen et al., 2021) in order to tailor welfare assessment to their unique patterns. These baselines can be established quite fast; Zambelis et al. (2019) showed that after 4 measurements of rising and lying down motions on one individual, the variability "flattened" and concluded that 4 measurements were a sufficient predictor of overall daily values. Sensors theoretically offer an opportunity to obtain these measurements on all individuals in a barn.

The different rates of changes in Figure 5 of Paper III, along with the significant baseline effect on the COM forward shift suggests not only that individuals have different patterns but that they respond differently, in this case to measures for improved welfare. By using sensors to look at the changes of the “worst-performers” we can see not only how measures improve overall welfare, but how it levels the field and brings the most at-risk animals closer to the most comfortable. This is to some extent speculation, but it aligns with evidence that flexible neck rails accommodated well for the diversity in cow sizes and movement patterns (Brouwers et al., 2025).

6.2 Implications

6.2.1 Improving cubicles through objective measures on posture transitions and 3D pose

One of the criteria for selecting indicators was notably how straightforward it would be to calculate them from 3D poses. These indicators do hold a degree of ambiguity. In Paper II we had found a modest correlation between rising duration and crawling (r=0.41) supporting its value in informing on abnormal motions (Blokhuis et al., 2013). Yet, evidence in earlier cubicle designs found no difference in the duration of the lying down motion when it was classified as abnormal or as normal (de Vries 1987 reviewed by Lidfors 1989). We similarly found that total duration of rising did not differ between the rigid and flexible cubicles, and that lying down duration only differed relative to the first baseline stage. This might only hold in specific contexts but does show that duration is not an “iceberg indicator”. The results in the PCA suggest that other rising patterns exist, that combine fast movement with abnormal motions. One of the strengths of 3D pose, and of other sensors for this purpose, is that it is able to simultaneously score a variety of indicators, in the time and space dimension (and their derivative) to provide a broader picture, and possibly to identify different clusters of strategies that cows use to cope with a restrictive cubicle.

Studies suggest about 0.9m of forward lunge space; that is unobstructed space in front of the cubicle to lunge the head forward (Cook, 2019). In the research farm where the data for this thesis was collected, such space was provided; 1.65m to be exact. Yet, backwards crawling – which was interpreted as cows attempting to increase the space available in front of them – was still observed in 51.2% of cases (Kroese et al., 2025). In another study, cows were found to hit the bar work in 0 to 25% of rising bouts depending on the cow (Zambelis et al., 2019) and in 70% of lying bouts overall. In a study providing 0.65m of lunge space, cows lunged to the side rather than the front in 35.5% of cases (Brouwers et al., 2023b). From these results, it seems that offering forward space is not sufficient for ensuring that all cows can lunge forward unhindered. Head and neck rail will interfere with the motions (Veissier et al., 2004), and the provision of forward space may not be a sufficiently reliable resource-based indicator of movement opportunities. In-line with Welfare Quality, animal-based measures offer direct insight as to the effect of the environment on the animals (Blokhuis et al., 2013).

Motion capture does accurately measure spatial use (Ceballos et al., 2004) but is impractical for implementation at scale. Pose estimation provides the same output as motion capture and measures the displacement of anatomical structures, albeit with different levels of accuracy (Lawin et al., 2023). The technology presented in this thesis holds the potential to be a practical method of asking the animals directly if they are able to use the space they need when getting up and lying down.

We had suggested in Paper III that actual lunge distance should rather be employed to quantify movement opportunities. Interestingly, we found no change in lunge distance between the metal bars and flexible rails configurations. There was a considerable change in forward head displacement when lying down however, when the metal bars were put back in place. It is possible, and would be aligned with the current consensus, that forward lunge ability is driven more by the space available in front of the cubicle (Cook, 2009), rather than by the metal elements. It is also likely that the head bar was already positioned at a permissive level. Regardless of the reason, we have found a significant increase in the forward movement of the COM. This measure, which is made possible by pose estimation brings measured indicators one step closer to the biomechanical drivers behind the head lunge, of which the forward displacement is a component visible to the human eye.

Limitations need to be acknowledged to clarify the scope of the results in providing evidence towards the aims of the thesis. First, in quasi-experiments, the absence of randomisation introduces a risk of uncontrolled confounders. This limitation is mitigated by the repeated measures design, where each individual serves as its own control. This is preferable for studies of biomechanics (Leclercq et al., 2024; Tijssen et al., 2021). The duration of the intervention was short (2 weeks) compared to the cows' previous experience with regular cubicles (since they were heifers), which were sustained if the cows visited other cubicles outside of the experimental area. This means that motions in the flexible cubicles may may still have been influenced by past and daily experiences in rigid cubicles. The changes we observe do align with a longitudinal study supporting the long-term effects of flexible cubicles (Brouwers et al., 2025).

Second, technical challenges in pose estimation, such as occasional missing poses, drift, and difficulties in labelling narrow or occluded anatomical regions add noise to the estimates. In the return to baseline, we obtained about half the sample size as in the other phases. This is most likely due to a shift in calibration that reduced the sensitivity of the 3D tracking. This in turn means that technical challenges remain for long-term monitoring.

6.2.2 Continuous monitoring at scale

The technology presented here represents an opportunity to increase the frequency of welfare assessment (at least the indicators it is capable of monitoring). If we are cautious about the accuracy of the measured indicators with 3D poses, we can still take a step to the bigger picture, which is the estimation of the cow's overall well-being, acknowledge the limitations of either method and ask: "is it better to measure something accurately 4 times a year, or to measure it less accurately every day?". Zambelis et al (2019) measured total daily values for comfort indicators (the average across all bouts in 24h), and how well different numbers of repetitions correlated with the mean 24h value. They showed that from 4 measures onwards, it flattened out (in other words, that 4 repetitions if a good enough approximation of daily average). This means that capturing all events in one day doesn't provide much value. The value of these tools might instead lie in the medium to long term rather than capturing as many events as possible in the short term. Berckmans (2017) described the purpose of PLF with the following: "*Farmers get a warning when something goes wrong in such a way that the PLF system brings them to the animal(s) that need their attention at that moment*." Although this quote is nearly 10 years old, it reveals a focus on negative aspects of welfare that is predominant. We could also argue that the purpose of PLF, exemplified by cubicle design, can be to elevate the lives of animals by continuously gathering physiological and behavioural indicators and seeing in which designs they respond in a way that is favourable to the animal, before something goes wrong.

Drifts in data generation can stem from changing environmental conditions, replacement in individual subjects (Moons et al., 2012), and drifts in the underlying process itself. Drifts can degrade model performance over time, making prediction of welfare indicators potentially less reliable over time (Vázquez-Diosdado et al., 2019). Models trained on data collected in one period often display reduced accuracy when applied to later data. Empirical studies show that it is harder to predict "the last 20%" of a time series using “the first 80%” than to predict a random 20%” (Sheridan, 2013). In my own developments, the methods to detect specific phases of the posture transition proposed in Paper II did not perform as well in Paper III, prompting me to annotate the events instead. This demonstrates that even when the same setups, facilities, and general conditions are maintained, repeatability can remain unexpectedly low. This is a common issue and a threat to the credibility of PLF (Tuyttens et al., 2022).

6.2.3 Monitoring welfare with sensors vs visual observations

What I noticed when conducting this work, and when talking to fellows researching similar topics, is that as we automate welfare assessment, the approach to welfare slightly changes.

One example is the type of cues. For example, in the Framework *Fråga Kon*, assessors are given a degree of discretion if qualifying the rising bout as abnormal. Sensors do not have access to these cues (or at least models are not trained to recognize them, notably because of the difficulty in establishing a ground truth for subjective assessments). What sensors do have is the capacity to merge different sources of behavioural and physiological data which together increase the robustness of welfare-based alerts (Do et al., 2020).

Most of the existing work on automated behaviour monitoring focuses narrowly on technical development, attempting to classify a few common postures or behaviours, without specific applications (Antognoli et al., 2025). Many of these developments are done on clean datasets, with homogenous data quality across examples and no occlusion, limiting real-world applicability (Menezes et al., 2024). The promise of sensors for welfare assessment might lie in their ability to continuous, and high-resolution monitoring integrated into the animal's environment without fatigue. To that end, much of the further work needed is the interpretation of long-term patterns.

Brouwers et al. (2023b) sought to detect abnormal rising and lying down movements using accelerometers and supervised learning. Their work lead to the creation of an R package for analysing rising and lying down movements (Simmler & Brouwers, 2024). When they attempted to automate the detection of sideways lunge using accelerometers, they only reached moderate accuracy (65%) (Brouwers et al.; 2023b). The authors impute this to a discrepancy between the way data was labelled (straight vs angled lunge) and the continuous nature of sensor data. There were many misclassifications on ambiguous edge cases. In my own results, I have found lunge angle to be continuous. It had a mean of 166.1° ± 0.5, a median of 168.9° and a skewness to the left by -1.3. Importantly with regards to labelling sideways lunge, there was no clear cut-off in the distribution which would have indicated straight versus angled lunge (Kroese et al., 2024). This distribution can be seen on Figure 9. While an observer might be able to define side lunge based on the observed angle combined with subtle behavioural cues, it remains impossible for a model to learn meaningful binary distinctions in a monomodal distribution.

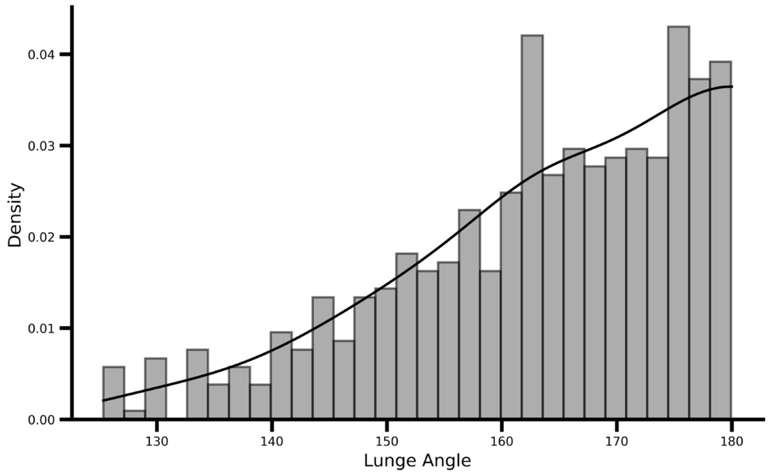


Figure 9. Distribution of lunge angles in rising events with kernel density estimation.

The main author of the aforementioned study on automating the detection of abnormal lunge, later made the statement "ethograms should be machine-learnable" (Brouwers et al., 2023). The system presented in this thesis applies this rationale, in the sense that we attempted to extract insights from the continuous data, without forcing labels. In that regard, we benefit from the high biomechanical interpretability of 3D pose estimation; if the cow lunges at an angle of, say, 153°, this is directly computable from the relative position of anatomical structures.

We mentioned earlier that the way the results were presented, reflected a certain normative stance, where improvements in welfare were sought "on average". Benchmarking and practical assessment methods are necessary for binding welfare regulations (Broom, 2017). Focusing on averages is a more robust method than taking, for example, the "worst-off" individuals, which might be subject to temporary ailments. These assessments are conducted by occasional inspections. With interest to sensor-based monitoring, we gain an opportunity for continuous monitoring of specific indicators. Thereby, we could get a more comprehensive timeline as to how each individual is faring, and if there are consistent clusters with impaired welfare.

6.3 Roadmap

6.3.1 More complex biomechanical modelling

The model used in this study was a simplified 2D, 2-rods 2-beams model. In our observations, we see that, since the cow is lying on the side, there is a lateral movement in the lift of the hind legs, suggesting that the model would benefit from incorporating the 3rd dimension. Introducing 3D joint rotations and capturing the full range of limb angles is expected to reduce errors generated by orthogonalization (Karashchuk et al., 2025). Differences between net static vertical forces and gravity show that the assumptions were imperfect.

In future developments, I hope to implement both 3 dimensions, and the reaction forces at the joints along the limbs. The existing set of keypoints were selected through consultation with experts in biomechanics, and correspond to true landmarks solicitated in posture transitions, providing a strong foundation for more complex modelling. This model would allow to understand how reaction forces propagate through the limb depending on joint angles. The introduction of supplementary articulations in the model (at the stifle and tarsi for example) would also increase the unknowns and address the current over-constraint.

By measuring the ground reaction forces (using loading cells placed under the hooves for example), we could produce a ground truth regarding weight distribution. We could "reverse engineer" the torque in each joint producing the observed trajectories (Özdil et al., 2025) and forces. Ultimately, this would allow a robust assessment of force propagation, and understanding of the constraints that the that the animal is able to sustain (Karashchuk et al., 2025). This would in turn be used to assess cubicles using pose estimation in 3D and force modelling, looking not only at the displacement of the centre of mass, but also the constraints placed on the joints, which could have a link with skin lesions.

6.3.2 Continuous monitoring with sensor fusion

Pose estimation in 3D applied to animals is a nascent field. The majority of the work focuses on technological development with testing on straightforward classification tasks like lying down versus standing (Antognoli et al., 2025). With this thesis, we contribute to developing the field one step further towards practical outcomes to improve the conditions of animals. Firstly, by proposing a data management and event detection framework that extracts information on data generated in near-real-time, getting closer to practical implementations. This framework is adapted to variations in the quality of the detections, although it does suffer from some false negative detections and limited sensitivity to missing detections. Secondly, by using the technology to generate meaningful information on the animal's comfort for behaviours where visual observations suffered with limitations in terms of scalability and quantifiability.

6.3.3 Expansion to other behaviours

The project set out to explore the place of multi-view pose estimation for welfare monitoring. Rising and lying down behaviours were one possible use case. Welfare assessment tools have tended to draw on a tradition inspired by the five freedoms in which the emphasis is on avoiding suffering (Broom, 2011). Increased attention has been brought to positive welfare (Rault et al., 2025), and on the need to develop technology in line with a multi-dimensional understanding of welfare (Foris et al., 2025). Behaviours indicative of positive welfare state such as social interactions or play are seldom used, and the difficulty in catching their occurrence might be one of the reasons. If we conceive sufficiently specific algorithms built on top of computer vision, or even of several sensor streams, we could catch these behaviours as they occur. In turn, the creation of longitudinal datasets from continuous monitoring can uncover patterns of occurrence, understanding the conditions under which these behaviours are displayed, and increase our knowledge of how they map to welfare.

One candidate behaviour in cattle is position of the ears. An observation study has linked different ear positions to feeding, brushing or queuing (D. De Oliveira & Keeling, 2018). Experimental results link ear position to alertness, stressors, and to positive experiences (Battini et al., 2019). Ear positioning remains contextual and ambiguous (Keeling et al., 2021), but continuous monitoring could support the validation of reliable indicators (Foris et al., 2025).

A different behaviour, likely inducive of positive states is brushing. Using the 2D poses from separate cameras from the multi-camera system, we were able to detect brushing bouts and identify which body segment a cow was brushing (Högberg et al., 2025). In turn, the objective is to develop a continuous ground truth, where we can link together the expression of specific behaviours (like brushing at the withers) and other experiences.

7. General conclusion

- Over half of all rising bouts, and over a quarter of lying down bouts in rigid cubicles were found to be indicative of compromised comfort, according to the thresholds found in existing literature. This highlights both the risks associated with cubicles, and specifically for this thesis, the ability for pose estimation in 3D to detect posture transitions and evaluate adverse welfare outcomes, at individual and at group levels.
- Flexible straps provide cows with a greater movement amplitude at the head lunge to get up, and potentially when lying down. The magnitude difference in lying down head displacement was comparable to that found in open packs.
- By monitoring each individual, we can tailor the assessment of specific welfare parameters to the variability and motion patterns of the individual rather than the herd average. Individuals with lower motion amplitudes disproportionately increased their rising motion highlighting how 3D pose could capture how the intervention can adapt to variability in the herd to "level the playing field".
- Features extracted from 3D poses estimation measure both the time and space dimension of rising bouts from which we can derive kinematic indicators, at high frequency. Using this technology provided novel quantitative information on the "bob room" and lunge angle, which we could not obtain with visual assessment yet is crucial to designing more comfortable cubicles.

References

Abade, C. C., Fregonesi, J. A., Von Keyserlingk, M. A. G., & Weary, D. M. (2015). Dairy cow preference and usage of an alternative freestall design. *Journal of Dairy Science*, *98*(2), 960–965. https://doi.org/10.3168/jds.2014-8527

Achour, B., Belkadi, M., Aoudjit, R., & Laghrouche, M. (2019). Unsupervised automated monitoring of dairy cows' behavior based on Inertial Measurement Unit attached to their back. *Computers and Electronics in Agriculture*, *167*, 105068. https://doi.org/10.1016/j.compag.2019.105068

Ådnegard Skarstad, G., Terragni, L., & Torjusen, H. (2007). Animal Welfare According to Norwegian Consumers and Producers: Definitions and Implications. *The International Journal of Sociology of Agriculture and Food*, 74-90 Pages. https://doi.org/10.48416/IJSAF.V15I3.285

Alonso, M. E., González-Montaña, J. R., & Lomillos, J. M. (2020). Consumers' Concerns and Perceptions of Farm Animal Welfare. *Animals*, *10*(3), 385. https://doi.org/10.3390/ani10030385

Antognoli, V., Presutti, L., Bovo, M., Torreggiani, D., & Tassinari, P. (2025). Computer Vision in Dairy Farm Management: A Literature Review of Current Applications and Future Perspectives. *Animals*, *15*(17), 2508. https://doi.org/10.3390/ani15172508

Arndt, S. S., Goerlich, V. C., & Van Der Staay, F. J. (2022). A dynamic concept of animal welfare: The role of appetitive and adverse internal and external factors and the animal's ability to adapt to them. *Frontiers in Animal Science*, *3*, 908513. https://doi.org/10.3389/fanim.2022.908513

Balasso, P., Marchesini, G., Ughelini, N., Serva, L., & Andrighetto, I. (2021). Machine Learning to Detect Posture and Behavior in Dairy Cows: Information from an Accelerometer on the Animal's Left Flank. *Animals*, *11*(10), 2972. https://doi.org/10.3390/ani11102972

Barkema, H. W., Von Keyserlingk, M. A. G., Kastelic, J. P., Lam, T. J. G. M., Luby, C., Roy, J.-P., LeBlanc, S. J., Keefe, G. P., & Kelton, D. F. (2015). Invited review: Changes in the dairy industry affecting dairy cattle health and welfare. *Journal of Dairy Science*, *98*(11), 7426–7445. https://doi.org/10.3168/jds.2015-9377

Barry, C., Ellingsen-Dalskau, K., Winckler, C., & Kielland, C. (2024). Exploring uses for an algorithmically generated Animal Welfare Indicator for welfare assessment of dairy herds. *Journal of Dairy Science*, *107*(6), 3941–3958. https://doi.org/10.3168/jds.2023-24158

Bastian, P., Basu, R., & Dette, H. (2024). Multiple change point detection in functional data with applications to biomechanical fatigue data. *The Annals of Applied Statistics*, *18*(4). https://doi.org/10.1214/24-AOAS1926

Battini, M., Agostini, A., & Mattiello, S. (2019). Understanding Cows' Emotions on Farm: Are Eye White and Ear Posture Reliable Indicators? *Animals*, *9*(8), 477. https://doi.org/10.3390/ani9080477

Beggs, D. S., Jongman, E. C., Hemsworth, P. H., & Fisher, A. D. (2019). The effects of herd size on the welfare of dairy cows in a pasture-based system using animal- and resource-based indicators. *Journal of Dairy Science*, *102*(4), 3406–3420. https://doi.org/10.3168/jds.2018-14850

Bernardi, F., Fregonesi, J., Winckler, C., Veira, D. M., von Keyserlingk, M. A. G., & Weary, D. M. (2009). The stall-design paradox: Neck rails increase lameness but improve udder and stall hygiene. *Journal of Dairy Science*, *92*(7), 3074–3080. https://doi.org/10.3168/JDS.2008-1166

Blokhuis, H., Miele, M., Veissier, I., & Jones, B. (2013). *Improving farm animal welfare* (H. Blokhuis, M. Miele, I. Veissier, & B. Jones, Eds). Brill | Wageningen Academic. https://doi.org/10.3920/978-90-8686-770-7

Bokkers, E., De Vries, M., Antonissen, I., & De Boer, I. (2012). Inter- and intra-observer reliability of experienced and inexperienced observers for the Qualitative Behaviour Assessment in dairy cattle. *Animal Welfare*, *21*(3), 307–318. https://doi.org/10.7120/09627286.21.3.307

Broom, D. M. (1996). Animal welfare defined in terms of attempts to cope with the environment. *Acta Agriculturae Scandinavica*, *27*, 22–28.

Broom, D. M. (2011). A History of Animal Welfare Science. *Acta Biotheoretica*, *59*(2), 121–137. https://doi.org/10.1007/s10441-011-9123-3

Broom, D. M. (2017). *Animal Welfare in the European Union* (Study No. PE 583.114). Directorate General for Internal Policies. https://www.europarl.europa.eu/RegData/etudes/STUD/2017/583114/IPOL_STU%282017%29583114_EN.pdf

Brouwers, S. P., Simmler, M., Savary, P., & Scriba, M. (2023a). Automatic detection of atypical head lunge movements of dairy cows in free-stall cubicles using accelerometers and machine learning. *Book of Abstracts*, 124.

Brouwers, S. P., Simmler, M., Savary, P., & Scriba, M. F. (2023b). Towards a novel method for detecting atypical lying down and standing up behaviors in dairy cows using accelerometers and machine learning. *Smart Agricultural Technology*, *4*, 100199. https://doi.org/10.1016/J.ATECH.2023.100199

Brouwers, S. P., Simmler, M., Scriba, M. F., & Savary, P. (2024). Cubicle design and dairy cow rising and lying down behaviours in free-stalls with insufficient lunge space. *Animal*, 101314. https://doi.org/10.1016/J.ANIMAL.2024.101314

Brouwers, S. P., Schug, A. F. E., Simmler, M., & Savary, P. (2025). The effect of neck strap positioning relative to dairy cow body size on rising, lying down, and defecation behaviour in lying cubicles. *Animal*, 101507. https://doi.org/10.1016/j.animal.2025.101507

Buller, H., Blokhuis, H., Lokhorst, K., Silberberg, M., & Veissier, I. (2020). Animal Welfare Management in a Digital World. *Animals*, *10*(10), 1779. https://doi.org/10.3390/ani10101779

Ceballos, A., Sanderson, D., Rushen, J., & Weary, D. M. (2004). Improving stall design: Use of 3-D kinematics to measure space use by dairy cows when lying down. *Journal of Dairy Science*, *87*(7), 2042–2050. https://doi.org/10.3168/JDS.S0022-0302(04)70022-3

Chapinal, N., Passillé, A. M. de, & Rushen, J. (2009). Weight distribution and gait in dairy cattle are affected by milking and late pregnancy. *Journal of Dairy Science*. https://doi.org/10.3168/jds.2008-1533

Cook, N. B. (2009). Free-stall design for maximum cow comfort. *WCDS Advances in Dairy Technology 21*, *21*, 255–268. https://acrre.ualberta.ca/wp-content/uploads/sites/57/wcds_archive/Archive/2009/Manuscripts/FreeStallDesign.pdf

Cook, N. B. (2019). Optimizing resting behavior in lactating dairy cows through freestall design. *Veterinary Clinics of North America: Food Animal Practice*, *35*(1), 93–109. https://doi.org/10.1016/J.CVFA.2018.10.005

Cook, N. B., Bennett, T. B., & Nordlund, K. V. (2005). Monitoring indices of cow comfort in free-stall-housed dairy herds. *Journal of Dairy Science*, *88*(11), 3876–3885. https://doi.org/10.3168/JDS.S0022-0302(05)73073-3

Cook, N. B., & Nordlund, K. (2005). Update on dairy cow free stall design. *The Bovine Practitioner*, *39*, 29–36. https://doi.org/10.21423/bovine-vol39no1p29-36

Cook, N. B., & Nordlund, K. V. (2009). The influence of the environment on dairy cow behavior, claw health and herd lameness dynamics. *The Veterinary Journal*, *179*(3), 360–369. https://doi.org/10.1016/j.tvjl.2007.09.016

De Oliveira, D., & Keeling, L. J. (2018). Routine activities and emotion in the life of dairy cows: Integrating body language into an affective state framework. *PLOS ONE*, *13*(5), e0195674. https://doi.org/10.1371/journal.pone.0195674

De Oliveira, F. M., Ferraz, G. A. E. S., André, A. L. G., Santana, L. S., Norton, T., & Ferraz, P. F. P. (2024). Digital and Precision Technologies in Dairy Cattle Farming: A Bibliometric Analysis. *Animals*, *14*(12), 1832. https://doi.org/10.3390/ani14121832

Dirksen, N., Gygax, L., Traulsen, I., Wechsler, B., & Burla, J.-B. (2020). Body size in relation to cubicle dimensions affects lying behavior and joint lesions in dairy cows. *Journal of Dairy Science*, *103*(10), 9407–9417. https://doi.org/10.3168/jds.2019-16464

Do, J. P., Defensor, E. B., Ichim, C. V., Lim, M. A., Mechanic, J. A., Rabe, M. D., & Schaevitz, L. R. (2020). Automated and Continuous Monitoring of Animal Welfare through Digital Alerting. *Comparative Medicine*, *70*(4), 313–327. https://doi.org/10.30802/AALAS-CM-19-000090

Fernandes, A. F. A., Dórea, J. R. R., & Rosa, G. J. D. M. (2020). Image Analysis and Computer Vision Applications in Animal Sciences: An Overview. *Frontiers in Veterinary Science*, *7*, 551269. https://doi.org/10.3389/fvets.2020.551269

Foris, B., Sheng, K., Dürnberger, C., Oczak, M., & Rault, J.-L. (2025a). AI for One Welfare: The role of animal welfare scientists in developing valid and ethical AI-based welfare assessment tools. *Frontiers in Veterinary Science*. https://doi.org/10.3389/fvets.2025.1645901

Foris, B., Sheng, K., Dürnberger, C., Oczak, M., & Rault, J.-L. (2025b). AI for One Welfare: The role of animal welfare scientists in developing valid and ethical AI-based welfare assessment tools. *Frontiers in Veterinary Science*, *12*, 1645901. https://doi.org/10.3389/fvets.2025.1645901

Ge, Z., Liu, S., Wang, F., Li, Z., & Sun, J. (2021). *YOLOX: Exceeding YOLO Series in 2021* (Version 2). arXiv. https://doi.org/10.48550/ARXIV.2107.08430

Gieseke, D., Lambertz, C., & Gauly, M. (2020). Effects of cubicle characteristics on animal welfare indicators in dairy cattle. *Animal*, *14*(9), 1934–1942. https://doi.org/10.1017/S1751731120000609

Gosztolai, A., Günel, S., Lobato-Ríos, V., Pietro Abrate, M., Morales, D., Rhodin, H., Fua, P., & Ramdya, P. (2021). LiftPose3D, a deep learning-based approach for transforming two-dimensional to three-dimensional poses in laboratory animals. *Nature Methods*, *18*(8), 975–981. https://doi.org/10.1038/s41592-021-01226-z

Gris, K. V., Coutu, J.-P., & Gris, D. (2017). Supervised and Unsupervised Learning Technology in the Study of Rodent Behavior. *Frontiers in Behavioral Neuroscience*, *11*, 141. https://doi.org/10.3389/fnbeh.2017.00141

Haley, D. B., de Passillé, A. M., & Rushen, J. (2001). Assessing cow comfort: Effects of two floor types and two tie stall designs on the behaviour of lactating dairy cows. *Applied Animal Behaviour Science*, *71*(2), 105–117. https://doi.org/10.1016/S0168-1591(00)00175-1

Haley, D. B., Rushen, J., & de Passillé, A. M. (2000). Behavioural indicators of cow comfort: Activity and resting behaviour of dairy cows in two types of housing. *Canadian Journal of Animal Science*, *80*(2), 257–263. https://doi.org/10.4141/A99-084

Hamäläinen, W., Järvinen, M., Martiskainen, P., & Mononen, J. (2011, November). Jerk-based feature extraction for robust activity recognition from acceleration data. *11th International Conference on Intelligent Systems Design and Applications*. https://doi.org/10.1109/ISDA.2011.6121760

Hansson, H., & Lagerkvist, C. J. (2015). Identifying use and non-use values of animal welfare: Evidence from Swedish dairy agriculture. *Food Policy*, *50*, 35–42. https://doi.org/10.1016/j.foodpol.2014.10.012

Hartley, R. I. (1997). In defense of the eight-point algorithm. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, *19*(6),

580–593. https://doi.org/10.1109/34.601246
Hartley, R., & Zisserman, A. (2003). *Multiple view geometry in computer vision*. Cambridge University Press.
Högberg, N., Berthet, D., Alam, M., Nielsen, P. P., Tamminen, L.-M., Fall, N., & Kroese, A. (2025). Exploring pose estimation as a tool for the assessment of brush use patterns in dairy cows. *Applied Animal Behaviour Science*, *292*, 106746. https://doi.org/10.1016/j.applanim.2025.106746
Högberg, N., Lidfors, L., Hessle, A., Arvidsson Segerkvist, K., Herlin, A., & Höglund, J. (2019). Effects of nematode parasitism on activity patterns in first-season grazing cattle. *Veterinary Parasitology*, *276*, 100011. https://doi.org/10.1016/j.vpoa.2019.100011
Huang, S., & Moliner, O. (2022). *Extrinsic calibration of multi-camera system* (United States Patent and Trademark Office (USPTO) Patent No. US 11,475,595 B2). https://patents.google.com/patent/US11475595B2/en
Ito, K., Von Keyserlingk, M. A. G., LeBlanc, S. J., & Weary, D. M. (2010). Lying behavior as an indicator of lameness in dairy cows. *Journal of Dairy Science*, *93*(8), 3553–3560. https://doi.org/10.3168/jds.2009-2951
Jang, D. H., Kim, C., Ko, Y.-G., & Kim, Y. H. (2020). Estimation of Body Weight for Korean Cattle Using Three-Dimensional Image. *Journal of Biosystems Engineering*, *45*(4), 325–332. https://doi.org/10.1007/s42853-020-00073-8
Johnson, W. R., Mian, A., Donnelly, C. J., Lloyd, D., & Alderson, J. (2018). Predicting athlete ground reaction forces and moments from motion capture. *Medical & Biological Engineering & Computing*, *56*(10), 1781–1792. https://doi.org/10.1007/s11517-018-1802-7
Karashchuk, L., Li, J. S., Chou, G. M., Walling-Bell, S., Brunton, S. L., Tuthill, J. C., & Brunton, B. W. (2025). Sensorimotor delays constrain robust locomotion in a 3D kinematic model of fly walking. *eLife*, *13*, RP99005. https://doi.org/10.7554/eLife.99005.3
Keeling, L. J., Winckler, C., Hintze, S., & Forkman, B. (2021). Towards a Positive Welfare Protocol for Cattle: A Critical Review of Indicators and Suggestion of How We Might Proceed. *Frontiers in Animal Science*, *2*, 753080. https://doi.org/10.3389/fanim.2021.753080
Kohli, E. (1987). Vergleich des Abliegeverhaltens von Milchkühen auf der Weide und im Anbindestall. *Aktuelle Arbeiten Zur Artgemässen Tierhaltung*, *319*, 18–38.
Kroese, A., Alam, M., Hernlund, E., Berthet, D., Tamminen, L.-M., Fall, N., & Högberg, N. (2024). 3D pose estimation to detect posture transition in free-stall housed dairy cows. *Journal of Dairy Science*, *107*(9). https://doi.org/10.3168/jds.2023-24427
Kroese, A., Högberg, N., Berthet, D., Tamminen, L.-M., Fall, N., & Alam, M. (2024). Exploring the link between cow size and sideways lunging using 3D

pose estimation. *11th European Conference on Precision Livestock Farming*, 32–39. https://www.diva-portal.org/smash/record.jsf?pid=diva2%3A1916993&dswid=7796

Kroese, A., Högberg, N., Diaz Vicuna, E., Berthet, D., Fall, N., Alam, M., & Tamminen, L.-M. (2025). Evaluating the automated measurement of abnormal rising and lying down behaviours in dairy cows using 3D pose estimation. *Smart Agricultural Technology*, *12*, 101205. https://doi.org/10.1016/j.atech.2025.101205

Krohn, C. C., & Munksgaard, L. (1993). Behaviour of dairy cows kept in extensive (loose housing/pasture) or intensive (tie stall) environments II. Lying and lying-down behaviour. *Applied Animal Behaviour Science*, *37*(1), 1–16. https://doi.org/10.1016/0168-1591(93)90066-X

Lawin, F. J., Byström, A., Roepstorff, C., Rhodin, M., Almlöf, M., Silva, M., Andersen, P. H., Kjellström, H., & Hernlund, E. (2023). Is markerless more or less? Comparing a smartphone computer vision method for equine lameness assessment to multi-camera motion capture. *Animals*, *13*(3), 390. https://doi.org/10.3390/ani13030390

Leclercq, A., Ask, K., Mellbin, Y., Byström, A., Bragança, F. M. S., Söderlind, M., Telezhenko, E., Bergsten, C., Andersen, P. H., Rhodin, M., & Hernlund, E. (2024). Kinematic changes in dairy cows with induced hindlimb lameness: Transferring methodology from the field of equine biomechanics. *Animal*. https://doi.org/10.1016/j.animal.2024.101269

Lee, J. G., Lee, S. S., Alam, M., Lee, S. M., Seong, H.-S., Park, M. N., Han, S., Nguyen, H.-P., Baek, M. K., Phan, A. T., Dang, C. G., & Nguyen, D. T. (2024). Utilizing 3D Point Cloud Technology with Deep Learning for Automated Measurement and Analysis of Dairy Cows. *Sensors*, *24*(3), 987. https://doi.org/10.3390/s24030987

Legrand, A. L., Von Keyserlingk, M. A. G., & Weary, D. M. (2009). Preference and usage of pasture versus free-stall housing by lactating dairy cattle. *Journal of Dairy Science*, *92*(8), 3651–3658. https://doi.org/10.3168/jds.2008-1733

Lidfors, L. (1989). The use of getting up and lying down movements in the evaluation of dairy cattle environments. *Veterinary Research Communications*, *13*(4), 307–324. https://doi.org/10.1007/BF00420838

Lindena, T., & Hess, S. (2022). Is animal welfare better on smaller dairy farms? Evidence from 3,085 dairy farms in Germany. *Journal of Dairy Science*, *105*(11), 8924–8945. https://doi.org/10.3168/jds.2022-21906

Linstädt, J., Thöne-Reineke, C., & Merle, R. (2024). Animal-based welfare indicators for dairy cows and their validity and practicality: A systematic review of the existing literature. *Frontiers in Veterinary Science*, *11*. https://doi.org/10.3389/fvets.2024.1429097

Liu, N., Qi, J., An, X., & Wang, Y. (2023). A Review on Information Technologies Applicable to Precision Dairy Farming: Focus on Behavior, Health Monitoring, and the Precise Feeding of Dairy Cows. *Agriculture*, *13*(10), 1858. https://doi.org/10.3390/agriculture13101858

Ma, S., Zhang, Q., Li, T., & Song, H. (2022). Basic motion behavior recognition of single dairy cow based on improved Rexnet 3D network. *Computers and Electronics in Agriculture*, *194*, 106772. https://doi.org/10.1016/j.compag.2022.106772

Marino, R., Petrera, F., & Abeni, F. (2023). Scientific Productions on Precision Livestock Farming: An Overview of the Evolution and Current State of Research Based on a Bibliometric Analysis. *Animals*, *13*(14), 2280. https://doi.org/10.3390/ani13142280

Maroto Molina, F., Pérez Marín, C. C., Molina Moreno, L., Agüera Buendía, E. I., & Pérez Marín, D. C. (2020). Welfare Quality® for dairy cows: Towards a sensor-based assessment. *Journal of Dairy Research*, *87*(S1), 28–33. https://doi.org/10.1017/S002202992000045X

Mellor, D. (2016). Updating Animal Welfare Thinking: Moving beyond the "Five Freedoms" towards "A Life Worth Living". *Animals*, *6*(3), 21. https://doi.org/10.3390/ani6030021

Mellor, D. J., Beausoleil, N. J., Littlewood, K. E., McLean, A. N., McGreevy, P. D., Jones, B., & Wilkins, C. (2020). The 2020 Five Domains Model: Including Human–Animal Interactions in Assessments of Animal Welfare. *Animals*, *10*(10), 1870. https://doi.org/10.3390/ani10101870

Menezes, G. L., Mazon, G., Ferreira, R. E. P., Cabrera, V. E., & Dorea, J. R. R. (2024). Artificial intelligence for livestock: A narrative review of the applications of computer vision systems and large language models for animal farming. *Animal Frontiers*, *14*(6), 42–53. https://doi.org/10.1093/af/vfae048

Moliner, O., Huang, S., & Astrom, K. (2021). Better prior knowledge improves human-pose-based extrinsic camera calibration. *2020 25th International Conference on Pattern Recognition (ICPR)*, 4758–4765. https://doi.org/10.1109/ICPR48806.2021.9411927

Moons, K. G. M., Kengne, A. P., Grobbee, D. E., Royston, P., Vergouwe, Y., Altman, D. G., & Woodward, M. (2012). Risk prediction models: II. External validation, model updating, and impact assessment. *Heart*, *98*(9), 691–698. https://doi.org/10.1136/heartjnl-2011-301247

Munksgaard, L., Jensen, M. B., Pedersen, L. J., Hansen, S. W., & Matthews, L. (2005). Quantifying behavioural priorities—Effects of time constraints on behaviour of dairy cows, Bos taurus. *Applied Animal Behaviour Science*, *92*(1–2), 3–14. https://doi.org/10.1016/J.APPLANIM.2004.11.005

Neveux, S., Weary, D. M., Rushen, J., Von Keyserlingk, M. A. G., & De Passillé, A. M. (2006). Hoof Discomfort Changes How Dairy Cattle Distribute Their

Body Weight. *Journal of Dairy Science*, *89*(7), 2503–2509. https://doi.org/10.3168/jds.S0022-0302(06)72325-6

Nielsen, S. S., Alvarez, J., Bicout, D. J., Calistri, P., Canali, E., Drewe, J. A., Garin-Bastuji, B., Rojas, J. L. G., Schmidt, C. G., Herskin, M., Michel, V., Chueca, M. Á. M., Padalino, B., Roberts, H. C., Spoolder, H., Stahl, K., Velarde, A., Viltrop, A., des Roches, A. D. B., … Winckler, C. (2023). Welfare of dairy cows. *EFSA Journal*, *21*(5). https://doi.org/10.2903/j.efsa.2023.7993

Nogueira, A. F. R., Oliveira, H. P., & Teixeira, L. F. (2025). Markerless multi-view 3D human pose estimation: A survey. *Image and Vision Computing*, *155*, 105437. https://doi.org/10.1016/j.imavis.2025.105437

Özdil, P. G., Ning, C., Phelps, J. S., Wang-Chen, S., Elisha, G., Blanke, A., Ijspeert, A., & Ramdya, P. (2025). *Musculoskeletal simulation of limb movement biomechanics in Drosophila melanogaster* (No. arXiv:2509.06426). arXiv. https://doi.org/10.48550/arXiv.2509.06426

Palczynski, L. (2019). *WP5 Community of Practice Deliverable 5.4—Third Annual Report for Researchers on Research Priorities on the Use of Sensor Technologies to Improve Productivity and Sustainability on Dairy Farms. 4D4FData Driven Dairy Decisions for Farmers Report; EU Horizon 2020 Grant Agreement No. 696367. 2019*.

Paranhos Da Costa, M. J. R., Taborda, P. A. B., De Lima Carvalhal, M. V., & Valente, T. S. (2021). Individual differences in the behavioral responsiveness of F1 Holstein-Gyr heifers to the training for milking routine. *Applied Animal Behaviour Science*, *241*, 105384. https://doi.org/10.1016/j.applanim.2021.105384

Qu, H., Zhang, H., Ban, Q., & Zhao, X. (2024). Accuracy and Adaptability Improvement in Aerobic Training: Integration of Self-Attention Mechanisms in 3D Pose Estimation and Kinematic Modeling. *IEEE Access*, *12*, 112470–112481. https://doi.org/10.1109/ACCESS.2024.3423765

Rault, J.-L., Bateson, M., Boissy, A., Forkman, B., Grinde, B., Gygax, L., Harfeld, J. L., Hintze, S., Keeling, L. J., Kostal, L., Lawrence, A. B., Mendl, M. T., Miele, M., Newberry, R. C., Sandøe, P., Špinka, M., Taylor, A. H., Webb, L. E., Whalin, L., & Jensen, M. B. (2025). A consensus on the definition of positive animal welfare. *Biology Letters*, *21*(1), 20240382. https://doi.org/10.1098/rsbl.2024.0382

Ren, K., Alam, M., Nielsen, P. P., Gussmann, M., & Rönnegård, L. (2022). Interpolation Methods to Improve Data Quality of Indoor Positioning Data for Dairy Cattle. *Frontiers in Animal Science*, *3*, 896666. https://doi.org/10.3389/fanim.2022.896666

Riaboff, L., Shalloo, L., Smeaton, A. F., Couvreur, S., Madouasse, A., & Keane, M. T. (2022). Predicting livestock behaviour using accelerometers: A systematic review of processing techniques for ruminant behaviour prediction from raw accelerometer data. *Computers and Electronics in Agriculture*, *192*, 106610. https://doi.org/10.1016/j.compag.2021.106610

Rutten, C. J., Velthuis, A. G. J., Steeneveld, W., & Hogeveen, H. (2013). Invited review: Sensors to support health management on dairy farms. *Journal of Dairy Science*, *96*(4), 1928–1952. https://doi.org/10.3168/jds.2012-6107

Sandøe, P., Corr, S., Lund, T., & Forkman, B. (2019). Aggregating animal welfare indicators: Can it be done in a transparent and ethically robust way? *Animal Welfare*, *28*(1), 67–76. https://doi.org/10.7120/09627286.28.1.067

Sapkota, S., Laven, R., Müller, K. R., & Kells, N. (2022). Practicability of a Time-Limited Welfare Assessment Protocol for Pasture-Based Dairy Farms, and a Preliminary Assessment of Welfare Outcome Thresholds. *Animals*, *12*(18), 2481. https://doi.org/10.3390/ani12182481

Schnitzer, U. (1971). Abliegen, Liegestellungen und Aufstehen beim Rind im Hinblick auf die Entwicklung von Stalleinrichtungen für Milchvieh. *Kuratorium Für Technik Und Bauwesen in Der Landwirtschaft-Bauschriften*, *10*(1), 43.

Sheridan, R. P. (2013). Time-Split Cross-Validation as a Method for Estimating the Goodness of Prospective Prediction. *Journal of Chemical Information and Modeling*, *53*(4), 783–790. https://doi.org/10.1021/ci400084k

Simmler, M., & Brouwers, S. P. (2024). *triact* package for R: Analyzing the lying behavior of cows from accelerometer data. *PeerJ*, *12*, e17036. https://doi.org/10.7717/peerj.17036

Smith, J. E., & Pinter-Wollman, N. (2021). Observing the unwatchable: Integrating automated sensing, naturalistic observations and animal social network analysis in the age of big data. *Journal of Animal Ecology*, *90*(1), 62–75. https://doi.org/10.1111/1365-2656.13362

Stygar, A. H., Gómez, Y., Berteselli, G. V., Dalla Costa, E., Canali, E., Niemi, J. K., Llonch, P., & Pastell, M. (2021). A Systematic Review on Commercially Available and Validated Sensor Technologies for Welfare Assessment of Dairy Cattle. *Frontiers in Veterinary Science*, *8*. https://doi.org/10.3389/fvets.2021.634338

Tamura, T., Okubo, Y., Deguchi, Y., Koshikawa, S., Takahashi, M., Chida, Y., & Okada, K. (2019). Dairy cattle behavior classifications based on decision tree learning using 3-axis neck-mounted accelerometers. *Animal Science Journal*, *90*(4), 589–596. https://doi.org/10.1111/asj.13184

Tangorra, F. M., Buoio, E., Calcante, A., Bassi, A., & Costa, A. (2024). Internet of Things (IoT): Sensors Application in Dairy Cattle Farming. *Animals*, *14*(21), 3071. https://doi.org/10.3390/ani14213071

Taylor, J. A. (2023). Precision agriculture. In *Encyclopedia of Soils in the Environment* (pp. 710–725). Elsevier. https://doi.org/10.1016/B978-0-12-822974-3.00261-5

Thompson, A. J., Weary, D. M., Bran, J. A., Daros, R. R., Hötzel, M. J., & Von Keyserlingk, M. A. G. (2019). Lameness and lying behavior in grazing dairy cows. *Journal of Dairy Science*, *102*(7), 6373–6382. https://doi.org/10.3168/jds.2018-15717

Tijssen, M., Serra Bragança, F. M., Ask, K., Rhodin, M., Andersen, P. H., Telezhenko, E., Bergsten, C., Nielen, M., & Hernlund, E. (2021). Kinematic gait characteristics of straight line walk in clinically sound dairy cows. *PLOS ONE*, *16*(7), e0253479. https://doi.org/10.1371/journal.pone.0253479

Truong, C., Oudre, L., & Vayatis, N. (2020). Selective review of offline change point detection methods. *Signal Processing*, *167*, 107299. https://doi.org/10.1016/J.SIGPRO.2019.107299

Tschanz, B., & Kämmer, P. (1979). The biology of behaviour as basis for the evaluation of resting areas of loose housing barns. *Tierzuechter*, *29*(4), 151–153.

Tucker, C. B., Jensen, M. B., Passillé, A. M. de, Hänninen, L., & Rushen, J. (2021). Invited review: Lying time and the welfare of dairy cows. *Journal of Dairy Science*, *104*(1), 20–46. https://doi.org/10.3168/JDS.2019-18074

Tucker, C. B., Zdanowicz, G., & Weary, D. M. (2006). Brisket Boards Reduce Freestall Use. *Journal of Dairy Science*, *89*(7), 2603–2607. https://doi.org/10.3168/jds.S0022-0302(06)72337-2

Tucker, C., Weary, D., & Fraser, D. (2004). Free-stall dimensions: Effects on preference and stall usage. *Journal of Dairy Science*, *87*(5), 1208–1216. https://doi.org/10.3168/jds.S0022-0302(04)73271-3

Tucker, C., Weary, D., Rushen, J., & de Passillé, A. M. (2004). Designing better environments for dairy cattle to rest. *Advances in Dairy Technology*, *16*, 39–53. https://wcds.ualberta.ca/wp-content/uploads/sites/57/wcds_archive/Archive/2004/Manuscripts/39Weary.pdf

Tuyttens, F. A. M., Molento, C. F. M., & Benaissa, S. (2022). Twelve Threats of Precision Livestock Farming (PLF) for Animal Welfare. *Frontiers in Veterinary Science*, *9*, 889623. https://doi.org/10.3389/fvets.2022.889623

Veissier, I., Capdeville, J., & Delval, E. (2004). Cubicle housing systems for cattle: Comfort of dairy cows depends on cubicle adjustment1. *Journal of Animal Science*, *82*(11), 3321–3337. https://doi.org/10.2527/2004.82113321x

Von Keyserlingk, M. A. G., Rushen, J., De Passillé, A. M., & Weary, D. M. (2009). Invited review: The welfare of dairy cattle—Key concepts and the role of science. *Journal of Dairy Science*, *92*(9), 4101–4111. https://doi.org/10.3168/jds.2009-2326

von Metzner, R. (1978). Analyse Tierischer Bewegungsabläufe zur Gestaltung artgemäßer Rinderkrippen. *Landtechnik*, *9*, 397–404.

Wang, J., Sun, K., Cheng, T., Jiang, B., Deng, C., Zhao, Y., Liu, D., Mu, Y., Tan, M., Wang, X., Liu, W., & Xiao, B. (2019). *Deep high-resolution representation learning for visual recognition* (Version 2). arXiv. https://doi.org/10.48550/ARXIV.1908.07919

Wegner, C. S., & Ternman, E. (2023). Lying behaviour of lactating dairy cows in a cow-calf contact freestall system. *Applied Animal Behaviour Science*, *259*, 105851. https://doi.org/10.1016/j.applanim.2023.105851

Wei, Y., Zhang, H., Gong, C., Wang, D., Ye, M., & Jia, Y. (2023). Study of Pose Estimation Based on Spatio-Temporal Characteristics of Cow Skeleton. *Agriculture*, *13*(8), 1535. https://doi.org/10.3390/agriculture13081535

Welfare Quality® Consortium. (2009). *Welfare Quality® Assessment Protocol for Cattle*. Welfare Quality® Consortium. https://www.welfarequalitynetwork.net/media/1088/cattle_protocol_without_veal_calves.pdf

Woodrum Setser, M. M., Neave, H. W., & Costa, J. H. C. (2024). Individuality of calves: Linking personality traits to feeding and activity daily patterns measured by precision livestock technology. *Journal of Dairy Science*, *107*(5), 3235–3251. https://doi.org/10.3168/jds.2023-24257

Ye, M., Xianwang Wang, Yang, R., Liu Ren, & Pollefeys, M. (2011). Accurate 3D pose estimation from a single depth image. *2011 International Conference on Computer Vision*,731–738. https://doi.org/10.1109/ICCV.2011.6126310

Zaffino Heyerhoff, J. C., LeBlanc, S. J., DeVries, T. J., Nash, C. G. R., Gibbons, J., Orsel, K., Barkema, H. W., Solano, L., Rushen, J., De Passillé, A. M., & Haley, D. B. (2014). Prevalence of and factors associated with hock, knee, and neck injuries on dairy cows in freestall housing in Canada. *Journal of Dairy Science*, *97*(1), 173–184. https://doi.org/10.3168/jds.2012-6367

Zambelis, A., Gagnon-Barbin, M., John, J. S., & Vasseur, E. (2019). Development of scoring systems for abnormal rising and lying down by dairy cattle, and their relationship with other welfare outcome measures. *Applied Animal Behaviour Science*, *220*. https://doi.org/10.1016/j.applanim.2019.104858

Popular science summary

How comfortable are cow when they go and lie down in modern dairy barns? The answer varies between farms but usually not as comfortable as they could be. Unlike pigs, cows do not defecate in specific “toilet areas” but rather wherever they are when they need to. Cows are also sensitive to udder infections that can be caused by contact with faeces. Do you see where this is going? If cows defecate on their bedding, there is a risk that they lie down in it and get infected. To prevent that, cows’ beds are delimited by cubicles; a rectangle made of metal bars on three sides that positions the cow with its rear over an alley, where an automatic scraper will remove the faeces. The problem with these metal bars, is that they restrict cows’ movements when getting up and lying down. You see, cows are heavy, so getting up requires a lot of effort. To help, they thrust their head forward, which puts less weight on the hind limbs, facilitating the lift. If there is a metal bar in the way of the head… bonk! This isn’t a nice feeling. Instead of metal bars, some farms have been experimenting with ratchet straps that offer a cushioning if the cow pushes against them. The results are promising, but how do we offer systematic scientific evidence that they indeed improve cows’ ability to get up and lie down. To answer this, we teamed up with Sony (yes, the ones behind the PlayStation). With their cutting-edge technology, we could automatically detect the position of cows in cubicles, and track the motion of their heads, limbs, and back with centimetre accuracy. Cows, and specific parts of their body were detected automatically on the frames of synchronized cameras. Then, the location of these body parts was triangulated to produce a position in 3D. Using this, we could accurately track how much cows moved in regular cubicles and cubicles with flexible straps. We found that cows had greater movement amplitudes in flexible cubicles, and even estimated the force borne on the rear limbs. The results were not straightforward but generally point towards the fact that cubicles with flexible elements are indeed more comfortable for getting up and lying down, and accommodate for a greater diversity of cows.

Populärvetenskaplig sammanfattning

Hur bekväma är kor när de lägger sig i moderna mjölkstallar? Svaret varierar mellan gårdar, men vanligtvis är det inte så bekvämt som det kunde vara. Till skillnad från grisar gör kor inte sina behov på specifika ”toalettplatser”, utan där de befinner sig när behovet uppstår. Kor är också känsliga för juverinfektioner, som kan orsakas av kontakt med avföring. Förstår du vart detta leder? Om korna gör sina behov på sin ströbädd finns det en risk att de lägger sig i det och blir infekterade. För att förhindra detta avgränsas kornas bäddar av bås, en rektangel gjord av metallstänger på tre sidor som placerar kon med bakdelen över en gång där en automatisk skrapa tar bort avföringen. Problemet med dessa metallstänger är att de begränsar kornas rörelser när de reser sig och lägger sig. Kor är tunga, så det kräver mycket kraft att resa sig. För att underlätta detta skjuter de huvudet framåt, vilket minskar belastningen på bakbenen och underlättar lyftet. Om det finns en metallstång i vägen för huvudet... bonk! Det är ingen trevlig känsla. Istället för metallstänger har vissa gårdar experimenterat med spännband, som ger lite mer flexibilitet om kon trycker mot dem. Resultaten är lovande, men hur kan vi systematiskt bevisa att de verkligen förbättrar kornas förmåga att resa sig och lägga sig? För att göra detta samarbetade vi med Sony (ja, de som ligger bakom PlayStation). Med deras banbrytande teknik kunde vi automatiskt detektera kornas position i bås och spåra rörelserna i deras huvuden, ben och rygg med centimeterprecision. Korna och specifika delar av deras kroppar detekterades automatiskt på synkroniserade kameror. Därefter triangulerades dessa kroppsdelars position för att skapa en position i 3D. Med hjälp av detta kunde vi noggrant spåra hur mycket korna rörde sig i vanliga bås och bås med flexibla spännband. Vi fann att korna hade större rörelseamplituder i flexibla bås och kunde till och med uppskatta kraften som belastade bakbenen. Resultaten var inte entydiga, men pekar generellt på att bås med flexibla element är bekvämare för att resa sig och lägga sig och passar en större mångfald av kor.

Acknowledgements

Immense gratitude goes to Sony and particularly to **David Berthet** and **Marc Ahlse**. It has been a fantastic and fruitful collaboration, and this work would not have been a fraction of what it is without your outstanding contribution, inputs and bright mood.

The best supervision team one could ever dream of; **Niclas Högberg**, **Lena-Mari Tamminen**, **Moudud Alam** and **Nils Fall** for your enthusiasm, for inspiring me by being incredible both as researchers and as individuals,
for believing in me, for always pushing me (but never too far), for presenting me with opportunities to become a better researcher and person, and for being insanely smart and equally nice people all around.

If I stated all the reasons that make me grateful to have **Claire Wegner** by my side, this book would become too heavy. You should really be a co-author on this thesis. Thanks for being a role model in compassion for animals and academic integrity, for getting me to both the start and end of this endeavour, for insightful conversations, and for being by my side this whole time.

My fabulous **colleagues** at the Unit for Epidemiology, past and present, and particularly **Karin Berggren**, for making me grateful to come to work every day.

My **friends**, here and far away, for filling each day with laughter and warmth, for the pleasant company and the hangovers, for being weightlifting and volleyball coaches, academic role models and absolute nutjobs. Most particularly **Homayoon, Valeriia, Pablo, Vincent, Ditsa Lydia, Victor, Lise, Julian, Lorena, Isa, Juliana, Miguel, Lex, Fernando, Markos & Lea, Anna & Jesper, Hector, Léonie & Antoine**.

Mochi and **Miso** for being cute and lying on the keyboard to remind me that they are at least as important as this thesis.

Researchers whom I have met along the way, who opened the doors to the world, **Ruã Daros**, **Lena Lidfors, Alex Guzhva** and conference besties **Barbara Pichlbauer**, **Ariana Negreiro** and **Stijn Brouwers.**

Ylva Mellbin, **Anna Leclercq** and **Elin Hernlund**, whom I perhaps solicited out of the blue, because your reputation follows you, and who have taken numerous hours of expertise and kindness to better this thesis.

The **folk from the PhD councils and SLUSS-DN**. It has been an amazingly rewarding time and I want to thank the crème de la crème, **Antonia Hartmann**, **Bradley Sparkes**, **Poorva Sundararajan** and **Emma Bromark**. I also thank **Ylva Hillbur** for your leadership, both with the mind and with the heart.

The **staff at Lövtsa** for the care to the animals and the life lessons you taught me, **Marc, Marcin, Morgan** and **Johanna.**

My **Parents**, my **brothers**, and **Grandmothers** who live too far away. Claire's parents **Gwen** and **Don** for making me feel part of the family.

I wholeheartedly thank the funders, for believing in this project and making this research possible. **FORMAS**, for funding the overall project throughout its duration. **Sony**, for providing contribution in-kind, numerous hours, along with invaluable brainpower and good mood. **Valborg Jacobssons Fund** and **Petra Lundbergs Foundation** for funding a cow location system suite that will be used in work following this thesis. The **Beijer Laboratory of Animal Science**. The F**oundation for Animal Protection of Kronoberg County** for supporting the flexible neck rail experiment. **SLU's fund for doctoral internationalisation**, the **Royal Society of Forestry and Agriculture** and **ÅForsk** for enabling dissemination of this project at international conferences.

I am perhaps part of the last cohort to have been told: *"Your thesis is not going to write itself!"*. In this line, I would like to thank the **developers of Mistral, ChatGPT and Perplexity** for providing powerful free-to-use LLMs. The same goes to all **developers** thanklessly maintaining open-source libraries that form the cornerstone of modern data infrastructure.

I

J. Dairy Sci. 107:6878–6887
https://doi.org/10.3168/jds.2023-24427

3-Dimensional pose estimation to detect posture transition in freestall-housed dairy cows

Adrien Kroese,[1]* Moudud Alam,[2] Elin Hernlund,[3] David Berthet,[4] Lena-Mari Tamminen,[1] Nils Fall,[1] and Niclas Högberg[1]
[1]Department of Clinical Sciences, Faculty of Veterinary Medicine and Animal Science, Swedish University of Agricultural Sciences, Uppsala, Sweden, 756 51
[2]School of Information and Engineering, Dalarna University, Borlänge, Sweden, 783 33
[3]Department of Anatomy, Physiology and Biochemistry, Faculty of Veterinary Medicine and Animal Science, Swedish University of Agricultural Sciences, Uppsala, Sweden, 756 51
[4]Sony Nordic, Lund, Sweden 221 88

ABSTRACT

Freestall comfort is reflected in various indicators, including the ability for dairy cattle to display unhindered posture transition movements in the cubicles. To ensure farm animal welfare, it is instrumental for the farm management to be able to continuously monitor occurrences of abnormal motions. Advances in computer vision have enabled accurate kinematic measurements in several fields, such as human, equine, and bovine biomechanics. An important step upstream to measuring displacement during posture transitions is determining that the behavior is accurately detected. In this study, we propose a framework for detecting lying-to-standing posture transitions from 3-dimensional (3D) pose estimation data. A multiview computer vision system recorded posture transitions between December 2021 and April 2022 in a Swedish stall housing 183 individual cows. The output data consisted of the 3D coordinates of specific anatomical landmarks. The sensitivity of posture transition detection was 88.2%, and precision reached 99.5%. In analyzing those transition movements, breakpoints detected the timestamp of onset of the rising motion, which was compared with that annotated by observers. Agreement between observers, measured by intraclass correlation, was 0.85 between 3 human observers and 0.81 when adding the automated detection. The intra-observer mean absolute difference in annotated timestamps ranged from 0.4 s to 0.7 s. The mean absolute difference between each observer and the automated detection ranged from 1.0 s to 1.3 s. We found a significant difference in annotated timestamps between all observer pairs, but not between the observers and the automated detection, leading to the conclusion that the automated detection does not introduce a distinct bias. We conclude that the model is able to accurately detect the phenomenon of interest and that it is equitable to an observer.
Key words: computer vision, animal welfare assessment, freestall cubicle, pose estimation

INTRODUCTION

All cubicles in a dairy barn are usually identical, but a natural variability exists both in animal size relative to the cubicle (Dirksen et al., 2020) and in individual motion patterns and locomotor activity (Shepley et al., 2020). A factor of stall comfort, which affects lesion prevalence and lying time, is the ease with which a cow is able to get up and down in the cubicle (Zambelis et al., 2019). Ease of movement during posture transition was highlighted as an evaluation criteria for stall quality in relation to cow comfort by Lidfors (1989), who noted that cows in cubicles were more regularly seen performing abnormal motions (such as sideways lunging or horse-like rising) than those on pasture. Ceballos et al. (2004) analyzed the kinematics of posture transitions and found that cows used less longitudinal space when rising in a cubicle than on an open pack. Given the evidence for the link between restrictive movements and signs of reduced welfare (Beaver et al., 2021), the quality of posture transitions is included as an indicator in welfare assessment schemes such as Welfare Quality (Blokhuis et al., 2013).

Assessing ease of posture transition per se, rather than through indirect signs of reduced comfort such as hock lesions (Dirksen et al., 2020) or reduced lying time (Shewbridge Carter et al., 2021), is more challenging, and practical objective methods are needed (Brouwers et al., 2023). Visual observations noting the occurrence of abnormal behaviors are commonplace in farm management and welfare assessment schemes. Alternatively,

Received November 13, 2023.
Accepted March 13, 2024.
*Corresponding author: adrien.kroese@slu.se

The list of standard abbreviations for JDS is available at adsa.org/jds-abbreviations-24. Nonstandard abbreviations are available in the Notes.

ease of movement can be assessed quantitatively by measuring the displacement of anatomical landmarks throughout bouts of posture transition (Ceballos et al., 2004). Drawbacks exist for both approaches. The visual method relies on time-consuming, sporadic human observations. Although Zambelis et al. (2019) found excellent agreement between observers (kappa of 0.93 for getting-up movement ease), a degree of subjectivity always exists in visual scoring of animal movements (Chaplin and Munksgaard, 2001; Vasseur, 2017). The acquisition of 3-dimensional (**3D**) kinematics data by Ceballos et al. (2004) relied on fitting motion-capture reflectors on cows, requiring lengthy preparation and exposure of the equipment to damage. These limitations might be a reason behind the low sample size (n = 5 cows with at least 2 bouts per cow) in the latter study.

Considering the variability in cow sizes and kinematic profiles and the need for objective methods to assess ease of movement, we propose a framework to detect lying-to-standing (**LTS**) posture transitions from 3D pose estimation data. As a step in validating the potential of this method, the aim of this study was to measure the performance of a feature extractor in detecting the onset of LTS posture transitions compared with the human eye.

MATERIALS AND METHODS

The study presented here was approved by the ethical committee Uppsala djurförsöksetiska nämnd under approval 5.8.18-13069/2021. The 3 Rs in animal research were considered when using existing video material, previously and noninvasively collected.

Location and Animals

Recordings were obtained at the Swedish Livestock Research Centre's dairy barn (Uppsala, Sweden). The herd comprises Swedish Holstein and Swedish Red cattle housed indoors with access to pasture 120 d a year, between May and September. Video was recorded on 30 separate days (midnight to midnight), sampled for convenience, between December 8, 2021, and April 28, 2022. Because the barn is lit at all times, recordings were obtained at all times of day. An average of 51 cows were present simultaneously in the pen, with individuals being added and removed throughout, for a total of 183 different individuals having visited the pen during the study period. A total of 7 RGB cameras (G3 Bullet, Ubiquiti) were placed around an area approximately one-quarter of the pen, located closest to the sorting gate to the milking robot, and oriented toward the rows of cubicles so that all cubicles in the study ward, including forward lunge room defined as the 60 cm beyond the head rail, were visible

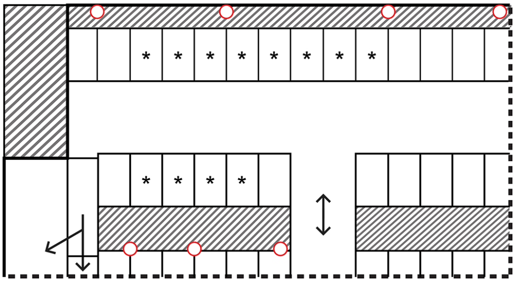


Figure 1. Schematic of the portion of the stall where recordings were obtained. The gray shaded areas are passageways unavailable to cows. Thick borders mark the stall boundaries, and dashed lines indicate a continuing area that is accessible to the cows beyond that shown here. Cameras are represented by red circles, placed between 2.8 and 3.6 m high. The parallel rectangles are cubicles; data were collected in cubicles marked with asterisks. The arrows indicate movement directions the cows are able to follow in the passageways.

by at least 2 cameras. The study ward comprised the 12 cubicles (CC1800 cubicle divider with rigid head bar, Delaval) for which video coverage was optimal, out of 66 total in the pen. The cameras were installed on fixed metal rails, part of the barn's infrastructure, between 2.8 and 3.6 m high. The locations of each camera, as well as the stall layout, are shown in Figure 1.

Cows had access to feeding troughs with ad-libitum mixed feed as well as 2 rotary brushes, and concentrate dispensed both at the milking robot and at concentrate dispensers. Passage through the milking robot's sorting gate was compulsory for access to the feed. Milking was done by one milking robot (VMS V300, Delaval), which cows had access to on a voluntary basis. Cows were brought to the robot by farm staff if they had not been milked in over 12 h.

Key Point Acquisition in 3 Dimensions

This study used 3D pose estimation software (Sony multi-camera system, Sony Nordic). The software estimates the 3D pose by finding cross-view correspondences across inferred 2-dimensional (**2D**) poses of the same object on synchronized views. It then creates a track for each object based on spatial continuity in the 3D location. The initial synchronization is achieved by reading the timestamp of each frame and relating the first full-second transition for a common timestamp across all video recordings as the initial synchronized frame. The initial frame synchronization is provided as an input to the multicamera system. Synchronization is maintained using the estimated time of arrival of each frame in the processing buffer.

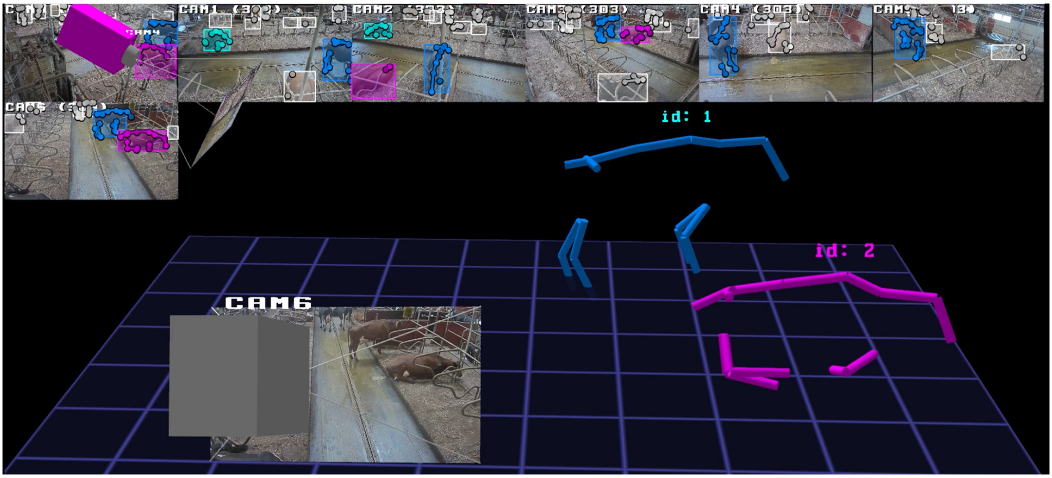


Figure 2. The 2D pose estimation and 3D fusion of 2 cows. The 2D results are displayed at the top, showing the synchronized frames from cameras 0 to 6, onto which predicted bounding boxes and key points are overlaid. The rest of the scene shows the projection of 2 cows from key points in 3D. Cameras 4 and 6 are represented as magenta and gray cuboids, respectively, in the 3D representation, in their spatial position relative to each other and to the cows. A projection of the frames from cameras 4 and 6 (identical to those in the 2D images above) is shown in front of the camera's 3D representation. The 5 other camera representations are not displayed from this angle, and camera 4 occludes the view from camera 0 because of the choice of angle. Only 4 of the key points shown in this figure were used in the study.

The 2D object detector and pose estimator use convolutional neural networks to detect cows and specific anatomical landmarks on RGB images, in the form of a bounding box and key points, respectively. The landmarks used in this study were limited to the center-top of the poll, the highest point at the withers, the spine at the 13th thoracic vertebra, and the top of the sacrum taken immediately behind the uppermost part of the ilium (referred to respectively as head, withers, t13, and sacrum).

The output data consists of one key point for each anatomical landmark with X, Y, and Z coordinates for each object and given frame. Figure 2 shows the estimated 3D position of the key points, linked to create a visual structure, for 2 objects during an LTS transition, as well as the video frames used to generate them.

Detection of Posture Transitions

The recordings were sampled visually by one observer with the aim of finding 1,000 sequences containing LTS transitions. When a cow was observed fully getting up from a lying positions, the timestamp was annotated, and a video sequence corresponding to a window of ± 15 s around the annotated timestamp was extracted. In the final data set, an arbitrary 979 sequences were eventually identified. These sequences were then processed with the 3D pose estimation software.

When the cow rises, the line formed by linking the sacrum and t13 key points increases its angle compared with the horizontal plane, as the cow's back is at an angle with the ground. By calculating the difference between the sacrum height and withers height, and following this difference through time, we identified peaks corresponding to LTS motions. When a peak above 0.4 (in the coordinates' arbitrary spatial reference system) was detected, the frame was considered to be within a potential rising motion. The mean withers Z position in the 120 frames located 330 frames after the peak was then compared with the mean withers Z position in the last 120 frames of the sequence. If the ratio of the height difference after and before the peak was higher than 140%, the track was classified as an LTS motion. Figure 3 illustrates this by showing the vertical position of the key points. At 16 s, there is an important difference in the heights of the withers (orange) and sacrum (green). This difference points toward a potential rising bout. Calculating the difference in withers position between the 5-s and 27-s marks, we determine that the animal has transitioned from a low, lying posture to a high, standing posture.

In these 979 sequences, this method initially detected 493 LTS motions for which the cow was tracked at each consecutive frame. For the remainder (486 sequences), the tracks were interrupted for several frames and the mo-

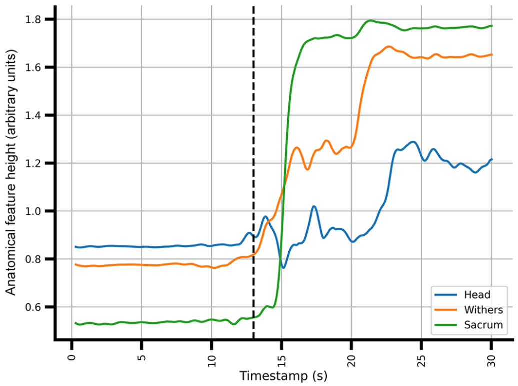


Figure 3. The coordinates of the anatomical landmarks of dairy cows were tracked with 3D pose estimation. This figure shows the Z coordinate (height) of a cow's head, withers, and sacrum throughout a lying-to-standing motion. Initially, the low variability on the vertical axis indicates that the cow is lying still. At about 11 s, the withers (orange) rise gently as the cow sits on its carps, followed by lunging with vertical bobbing of the head (blue) from 12 to 17 s. The sacrum (green) rises rapidly soon after, describing a sigmoid. There is a pause on the carps, with the sacrum already up, from 16 to 20 s. The cow has risen by the 22-s mark. The vertical dotted line shows the onset of the posture transition detected using linearly penalized segmentation. This example was selected for clarity.

tion was captured in several separate tracks. Detections were stitched together if they fit the following criteria:

- The tracks are found in the same 30-s sequence.
- The second track starts after the first track vanishes, and within an interval of 30 frames.
- The Euclidian distance in the 3D pose estimator's coordinate system between the last point in the vanishing track and the first point of the starting one is lower than 0.2.

No limit was imposed on the number of tracks appended together to form one single track, as long as the above conditions were fulfilled. The resulting stitched track was kept if it contained more than 700 frames, and discarded otherwise.

Using this method, an additional 370 rising sequences were detected by applying the height difference rule to the stitched tracks, giving a total of 863 predicted positives. For the remaining 116 sequences, either the animal was not detected by the pose estimation software, the posture transition detector failed to identify the occurrence, or the motion was split between different tracks that were not relatable due to noise or an interruption across more than 30 frames. Visual inspection of the predicted LTS motions revealed 4 false positives. In addition, 22 true positives were discarded from the data set because the posture transition was initiated before the start of the video snippet and thus not captured in its entirety.

Signal Processing

Each series of raw coordinates was processed to attenuate noise. A low pass filter with a cutoff frequency of 10 Hz was applied to remove high-frequency noise resulting from key point jittering. This cutoff was chosen based on the recommendations by Hamäläinen et al. (2011) and Riaboff et al. (2020) for noise removal on animal activity data. The filter was applied separately to each key point and the respective time series of its X, Y, and Z coordinates. The filter was implemented in Python 3.9 (Python Software Foundation) using the function "butter" from the SciPy package (Virtanen et al., 2020). Figure 3 illustrates the filtered Z coordinates time series during a rising sequence.

From the processed signal, consisting of the coordinates of each key point in 3 dimensions, we detected the timestamp at which the cow starts rising. Considering solely the kinematic features available through the 4 key points, this is most clearly reflected by the change in the position of the withers, as rising on the elbows will cause the withers to rise upward slightly, which is visible by an increase in the withers' Z (vertical) coordinate. When doing so, the cow aligns its back along the length of the cubicle, which is reflected in a change of the withers' Y coordinate (axis perpendicular to the cubicle's length). Although, from a behavior perspective, there is more to the LTS transition than solely the withers' movement, the system was blind to all but the position of 4 anatomical landmarks. The withers were chosen for the stability of the key-point (low jittering) and for their consistent motion pattern in the LTS transition across sequences. To detect the exact onset of rising motions, we used linearly penalized segmentation (Pelt), implemented the Python library "Ruptures" (Truong et al., 2020). Pelt was applied to the bivariate series of the Y (lateral, perpendicular to the cubicles) and Z (height) positions of the withers to identify breakpoints in the time series. No restrictions were set on the number of breakpoints to be detected. A baseline height (Z coordinate) was calculated for each sequence as the median withers height in the first 30 frames of the sequence. The break points detected by Pelt were iterated through. If the median withers height in the 30 frames following the breakpoint was higher than the baseline, the breakpoint was then considered to be the start of the rising motion. If not, we iterated to the next breakpoint and applied the same logic.

Data processing, feature extraction, and analyses were carried out in Python 3.9.3 using the packages NumPy 1.21.5 (Harris et al., 2020) and SciPy 1.9.1 (Virtanen et al., 2020).

Validation Experiment

To evaluate the performance of the tool in detecting the occurrence of LTS bouts, we compared the timestamps automatically detected to those annotated by 3 human observers, considered as the gold standard for behavioral observations. Observers were provided with the following definition: "The cow is lying down and rises on its breastbone and elbows, which causes the withers to rise visibly above the rest of the back." This definition is based on that of Lidfors (1989), but it adds the position of the withers as an indicator. The animals were seen to initiate the movement by centering their elbows under the body, this in turn causes the withers to rise slightly. This motion of the withers was used to determine the exact onset of the rising motion. The description was accompanied by illustrations taken from Schnitzer (1971) and Cermak (1988), as well as an ethogram describing the sequence of movements in the LTS transition, in which the movement to label was explicitly identified. This ethogram described the stages of the posture transition based on Lidfors (1989) and on Schnitzer (1971). Observers all received the same training, in which the ethogram was explained and examples were showcased; they reviewed 5 videos of different cows rising and agreed on the exact frame to label as the onset of the rising motion. These 5 videos were taken from the original data set and used solely for training the observers.

The validation data set was sampled randomly from the 471 complete LTS sequences captured in a single track. In total, 60 unique LTS sequences were annotated by at least 1 observer. This number was determined a priori, as no prior data were available on observer variability in posture transition detection. These sequences were the original 30 s synchronized video snips from which the key points were detected. The video was available to the observers from all 7 cameras used for key point detection, plus one additional ceiling mounted camera. Observers were free to choose the camera offering the best view of the animal performing the bout. Every observer was provided with a total of 55 randomly selected video clips. Of these 55 sequences, 30 were common to all observers and 10 were unique to each observer (40 different sequences per observer). The remaining 15 sequences were randomly resampled from the prior 40 and re-annotated by the same observer, to measure intra-observer reliability. All sequences were blinded, with a different label each time the sequence appeared.

Statistical Analysis

The mean absolute difference (**MAD**) in annotate timestamp was calculated between each observer to quantify intra-observer reliability as $MAD(i) = \frac{1}{15}\sum_{s_i=1}^{15}\left|\Delta_{s_i}\right|$, where $i = 1, 2,$ or 3 (observers) and $\Delta_{s_i} = t_{s_i,1} - t_{s_i,2}$, with $t_{s_i,1}$ and $t_{s_i,2}$ being the time stamp of the sth sequence provided at first and second assessment occasion, respectively, by observer i. In addition, the inter-rater MAD was calculated as $MAD(i,j) = \frac{1}{30}\sum_{s=1}^{15}\left|\Delta_{s,(i,j)}\right|$, where $\Delta_{s,(i,j)} = \left| t_{s,i} - t_{s,j} \right|, \forall i \neq j = 1,2,3$, with $t_{s,i}$ being the time stamp of the sth common sequence by ith observer. Mean differences (**MD**), as, for example, $MD(i) = \frac{1}{15}\sum_{s_i=1}^{15}\Delta_{s_i}$, were calculated, and the normality of Δ_{s_i} and $\Delta_{s,(i,j)}$ was assessed visually on a q-q plot. Subscripts i and j refer to 2 distinct observers: $i = 1, 2,$ or 3; $j = 2$ or 3. The MD and MAD indicate interobserver systematic bias and dispersion, respectively.

The following mixed effects models were fitted using statsmodels.formula.api.mixedlm (Seabold and Perktold, 2010) in Python 3.9 to evaluate the observer effect and intraclass correlation (ICC) with or without the automated detection:

$$t_{s,i,r} = \beta_0 + \beta_1 I(i=2) + \beta_2 I(i=3) + u_s + \varepsilon_{s,i,r}, \quad [1]$$

$$t_{s,i,r} = \beta_0 + \beta_1 I(i=2) + \beta_2 I(i=3) + \beta_3 M + u_s + \varepsilon_{s,i,r}, \quad [2]$$

where β_0 is the (fixed) intercept, $u_s \sim N(0,\sigma_u^2)$ is a random sequence effect, $s = 1$ to 40 is the sequence indicator, β_1 and β_2 are fixed observer effects, β_3 is a fixed effect corresponding to the automated detection taken as an additional observer (referred to as the "model" or M), and $\varepsilon_{s,i,r} \sim N(0,\sigma_e^2)$ is a (random) error term. The sequence number is indicated by the subscript s, I_i are the observers, and $r = 1, 2$ is the index for repeated sequences annotated 1 to 2 times by the same observer. The observer effects were tested using ANOVA. The ICC as a measure of interobserver agreement were calculated as $ICC = \frac{\sigma_u^2}{\sigma_u^2 + \sigma_e^2}$. A post hoc pairwise t-test with Bonferroni correction for 6 tests was then computed to test the pairwise differences between observers. The annotated timestamps were not normalized because a 1 s difference between observers, for example, has the same practical meaning in this context regardless of whether the annotation is done at the 4-s mark or the 12-s mark.

The performance of the algorithm was assessed in the same way, by treating the algorithm as an additional observer and seeing if it differed from the human observers. The differences were calculated between the algorithms' detection (denoted T^M) and the observer annotation, T^H. Bland-Altman plots were prepared for each observer pair $\left(T_i, T_j\right) = \left(\left\{t_{s,i}\right\}, \left\{t_{s,j}\right\}\right)$, and also comparing T^H with T^M, with a view to checking for the absence of a pattern and points beyond 1.96 standard deviations. Lastly, MAD(H, M), and MD(H, M) were calculated.

RESULTS

A total of 836 rising bouts were detected out of 979 visually selected sequences equating to a sensitivity of 88.5% or a false negative rate of 11.5%. Four sequences were wrongly classified as rising motions giving a precision of 99.5% or false positive rate of 0.5%.

Model 1, comparing only human observers, gave ICC = 0.85. We found a significant observer effect in predicting the annotated timestamps of LTS onset ($P < 0.001$) according to the ANOVA. When the model 2 was fitted to assess performance of the prediction, the ICC decreased to 0.81, remaining at a similarly satisfactory level of agreement. However, we found no significant difference between the predicted timestamp ("model") and each observer's annotations according to the post hoc pairwise *t*-test with Bonferroni correction of the type-1 error at $\alpha = 0.0083$. We identified a significant difference between all observer pairs: $P\ (T_1,T_2) = 0.0016$; $P\ (T_1,T_3) < 0.001$; $P\ (T_1,T_3) = 0.0018$.

Mean absolute differences $T_s^{M,H}$ are summarized in Table 1. These values indicate good interobserver agreement and good agreement between humans and machine. The magnitude of $T_s^{M,H}$ is identical to that of $T_s^{M,M}$, meaning that T^M could be used in further research, as the model does not deviate from the observers more than they do from one another. Figure 5 shows the timestamp annotated by each observer (including the model and repeat sequences) for each sequence.

Intra-observer reliability was assessed using the mean absolute difference in seconds, and consistency using the standard deviation (σ). Observer 1 had an MAD of 0.55 ± 0.88 s ($\mu \pm \sigma$). Observer 2 had an MAD of 0.68 ± 1.47 s, and observer 3 had an MAD of 0.36 ± 0.48 s. The pooled standard error was 0.27 s. The standard deviation is preferred here to the standard error to quantify the variability in the differences between and within observers in annotated timestamps, independently of the number of samples. These results indicate very good intra-observer reliability (under 1 s on average).

Finally, we compared the annotations to the automated detections visually using the Bland-Altman plot in Figure 4. The upper left plot shows most points to be centered around 0, without signs of consistent bias from the model. More importantly, the spread was similar when comparing observers to the algorithm and observers together.

Table 1. Interobserver agreement (MAD ± σ) between the annotations of all pairs of observers, including the model; pairs between observers calculate the MAD on 30 sequences, whereas pairs with the model include an additional 10 annotations, unique to each observer

Item	Observer 1	Observer 2	Observer 3
Model	1.02 ± 1.41	1.00 ± 1.70	1.30 ± 1.45
Observer 1		1.10 ± 1.26	1.67 ± 1.72
Observer 2			0.89 ± 1.01

DISCUSSION

The ICC values show a good agreement between automated model detection and human observers in detecting the onset of cows' rising motions, according to previous research on the use of ICC as a reliability metric in animal motion scoring (Kaler et al., 2009). The ANOVA demonstrated a significant observer effect, strengthening the claim that observations of cows' movements are prone to individual variations. The post hoc test showed a significant difference in annotated timestamps between all pairs of observers, but the difference between the model and the observers was not significant. We conclude from this that the model's detection lies somewhere in between the observers' annotations. The MD of −0.06 s between observers and the model (Figure 4) and the proximity of the points to 0 show that no systematic bias was introduced by the automated detection. This latter finding is also supported by Figure 5, showing the timestamp annotated by each observer at each sequence, in which there is no evidence of the detection being consistently divergent from human annotations, as the triangular points (model) are not systematically above or below the circular ones (observers). We also see that the predictions do not tend to be further from the annotations than the annotations are from each other.

This agreement is a crucial step in validating the capability of 3D computer vision to accurately identify this specific kinematic feature in bovine behavior. Notably, the findings suggest that the model's performance does not considerably differ from human observers when compared with the variability among human observers. This suggests that the model does not introduce a distinct source of error in the detection process. Although discrepancies exist between the model and human observations, the magnitude of these divergences is not meaningful in comparison to the overall duration of the LTS transition.

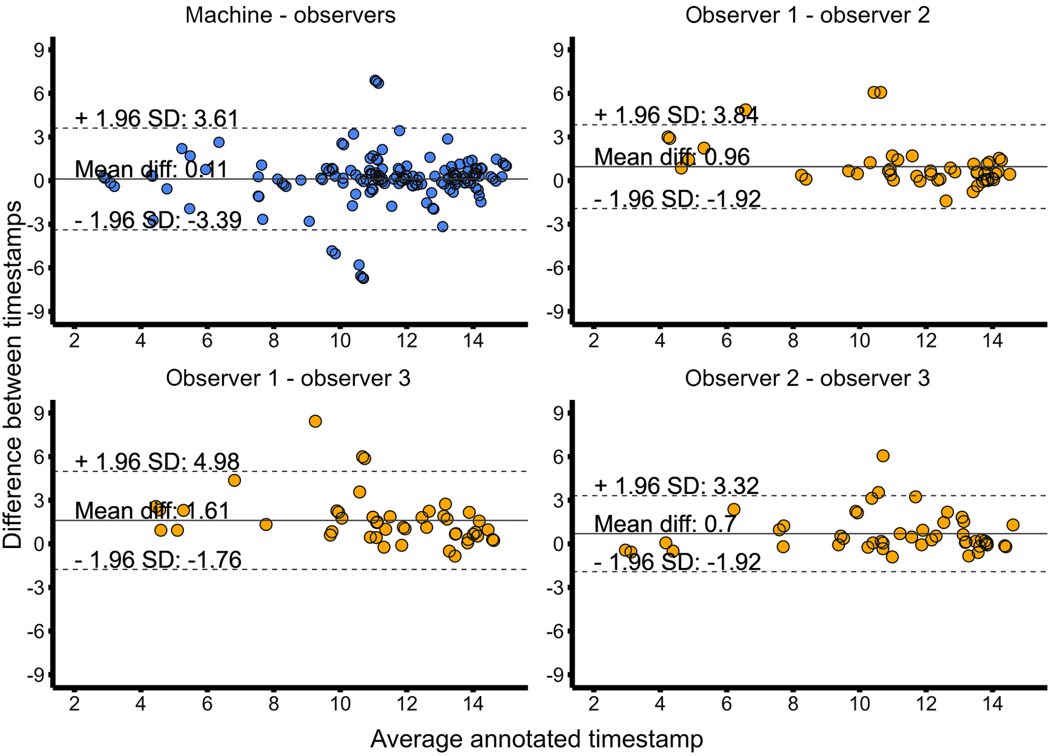


Figure 4. Bland-Altman plots comparing the timestamp of onset of cows' rising motions annotated by human observers to that predicted by the model. The 3D pose estimation provided the coordinates of cows' anatomical landmarks. Detecting breakpoints in the key point motion enabled detection of the onset of rising. Diff = difference. All units in seconds.

However, some limitations are important to mention. One such limitation is the likely over-representation of specific individuals. The animals were filmed in a limited area of the barn, and we can expect a degree of site fidelity from the animal (Vázquez Diosdado et al., 2018), leading to some individuals being over-represented. Because there was no individual detection, correcting for individuals was not possible. It is also unlikely that all recorded bouts were spontaneous; some may have been triggered by human intervention or by the presence of agonistic individuals. Bout motivation could introduce changes in kinematic patterns and velocity and potentially affect the accuracy of the automated detection.

Limitations also exist regarding external validity, as the study was conducted in a single cubicle design, under a limited period of time, and using manually selected video sequences. This manual selection work upstream of the automated processing is an important limitation that drove the high sensitivity and specificity. The same system should be tested on continuous recordings. To counterbalance this limitation, however, the posture transition is an evident behavior, with a large difference in key point height before and after, which would easily be captured even with noisy key points by simply following the height of the cow's back.

The scope of this study was determined retrospectively; the decision to compare the automated detection to manual annotations was made after collecting the data and visually identifying LTS motions. The inclusion criteria were based on data quality and not experimental considerations. The exclusion of 22 longer bouts discarded important information with implications for the most vulnerable individuals when it comes to stall comfort, as a long pause during the posture transition is associated with adverse welfare outcomes (Zambelis et al., 2019)

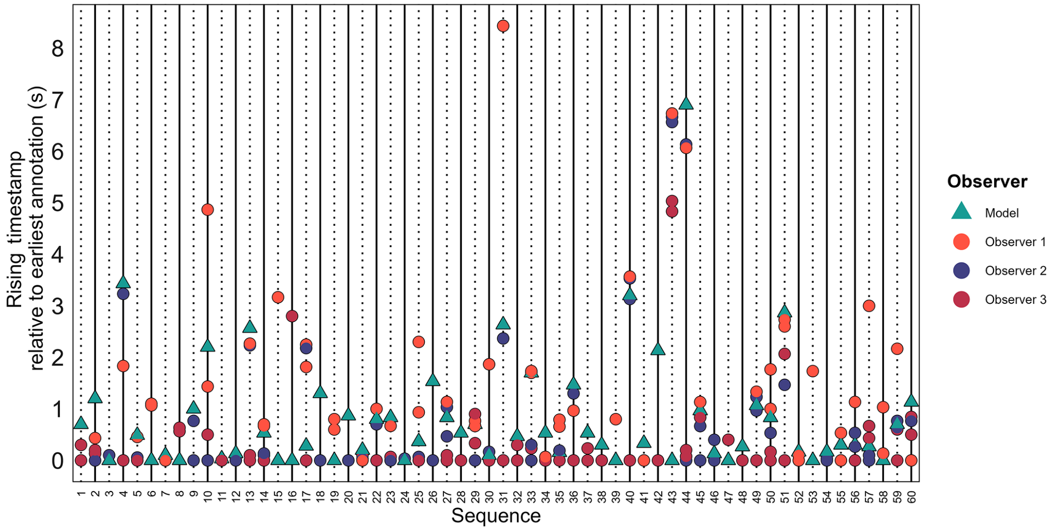


Figure 5. Annotated timestamp by each observer and by the model. The discrete x-axis shows each lying-to-standing sequences. On the y-axis is the timestamp of the onset of the posture transition annotated by each observer or predicted. From each annotation is subtracted the earliest timestamp in that sequence.

The study's gold standard was human observation, which is known to be variable across observers due to individual subjectivity. Although a bias is incorporated in the model, this bias is consistent across observations. The accuracy of the model could be improved by both altering the ethograms to make them more "machine-learnable" (Brouwers et al., 2023) and by diversifying the data. Importantly, although human observations are biased, humans are rarely completely incorrect, especially when the phenomenon at hand, such as posture transition, is evident. Algorithms on the other hand sometimes produce unexpected results, and monitoring and understanding their occurrence is essential for practical application. For instance, a difference of 6 s is found between the model and observer 2 in sequence 31 (Figure 5). Upon visual inspection of this sequence, the algorithm picked up on the onset of the adjustment movements, which were particularly lengthy in this sequence, making up the initial part of the posture transition. The second observer, on the other hand, noted the moment the fast rising motion occurred. This is not an error of either method, but a misalignment in the interpretation of the behavior. Referring to the description of the behavior provided to the observers, and quoted in the Materials and Methods section, the timestamp automatically detected is closer to the phenomenon of interest.

Most significant for this research is that automated detection via computer vision offers an objective method for detecting specific motions, which is desirable for studies of behavior and motion patterns. Judging by the advances in equine kinematic research, markerless computer vision constitutes both a robust and practical data acquisition tool to measure the displacement of anatomical landmarks, offering similar accuracy to motion capture, albeit for specific motions (Lawin et al., 2023). Reliably identifying the motion of interest is only a step in the study of posture transition kinematics, which contain welfare indicators (Zambelis et al., 2019), the measure of which can be automated (Brouwers et al., 2023). Future studies using this technology aim at implementing individual recognition, which could contribute to a pool of sensor data at individual level. However, in the absence of individual identification, this technology is still able to deliver meaningful information either at herd or at cubicle level. The automated detection through 3D computer vision could, after further validation, serve as a new gold standard for the task of detecting LTS transitions (and other movements), similar to how interpreting accelerometer data has become standard in behavior classification of ruminants (Riaboff et al., 2022).

CONCLUSIONS

In summary, our results demonstrate good agreement between human observations and automated detection of cows' rising motions. Notably, they indicate that the model introduces no more bias than human observers. This finding validates the use of multiview 3D pose estimation for detecting the onset of rising motions in bovine behavior, albeit in the conditions of a single farm. Automating the task with computer vision presents an opportunity to scale up bovine kinematic measurements and behavior monitoring and apply objective methods to further study.

NOTES

The authors thank the Swedish Research Council Formas (Stockholm, Sweden) for funding this research, the personnel of the Swedish Dairy Research Center (Uppsala, Lövsta, Sweden) for their outstanding support, and Sony Nordic (Lund, Sweden) for their extensive collaboration. The study presented here was approved by the ethical committee Uppsala djurförsöksetiska nämnd under approval 5.8.18-13069/2021. The authors declare that Sony Nordic has contributed to this research in kind and in staff hours. Sony Nordic provided the technology to generate the 3D pose. Conceptualization, study design, statistical analysis and presentation of the results were decided by researchers at the Swedish University of Agricultural Sciences (Uppsala, Sweden). Sony Nordic contributed in drafting the methods section regarding key-point acquisition in 3 dimensions, and in revising the final manuscript. This study was not conducted with the purpose of supporting a commercial claim. The authors have not stated any other conflicts of interest.

Nonstandard abbreviations used: 2D = 2-dimensional; 3D = 3-dimensional; diff = difference; ICC = intraclass correlation; LTS = lying-to-standing; MAD = mean absolute difference; MD = mean difference.

REFERENCES

Beaver, A., D. M. Weary, and M. A. G. von Keyserlingk. 2021. Invited review: The welfare of dairy cattle housed in tiestalls compared to less-restrictive housing types:—A systematic review. J. Dairy Sci. 104:9383–9417. https://doi.org/10.3168/jds.2020-19609.

Blokhuis, H., M. Miele, I. Veissier, and B. Jones. 2013. Improving Farm Animal Welfare: Science and Society Working Together—The Welfare Quality Approach. Wageningen Academic Publishers.

Brouwers, S. P., M. Simmler, P. Savary, and M. F. Scriba. 2023. Towards a novel method for detecting atypical lying down and standing up behaviors in dairy cows using accelerometers and machine learning. Smart Agric. Technol. 4:100199. https://doi.org/10.1016/j.atech.2023.100199.

Ceballos, A., D. Sanderson, J. Rushen, and D. M. Weary. 2004. Improving stall design: Use of 3-d kinematics to measure space use by dairy cows when lying down. J. Dairy Sci. 87:2042–2050. https://doi.org/10.3168/jds.S0022-0302(04)70022-3.

Cermak, J. 1988. Cow comfort and lameness: Design of cubicles. Bov. Pract. (Stillwater) 23: 79–83. https://doi.org/10.21423/bovine-vol0no23p79-83.

Chaplin, S., and L. Munksgaard. 2001. Evaluation of a simple method for assessment of rising behaviour in tethered dairy cows. Anim. Sci. 72:191–197. https://doi.org/10.1017/S1357729800055685.

Dirksen, N., L. Gygax, I. Traulsen, B. Wechsler, and J.-B. Burla. 2020. Body size in relation to cubicle dimensions affects lying behavior and joint lesions in dairy cows. J. Dairy Sci. 103:9407–9417. https://doi.org/10.3168/jds.2019-16464.

Hamäläinen, W., M. Järvinen, P. Martiskainen, and J. Mononen. 2011. Jerk-based feature extraction for robust activity recognition from acceleration data. 11th Int. Conf. on Intelligent Systems Design and Applications. IEEE. https://doi.org/10.1109/ISDA.2011.6121760.

Harris, C. R., K. J. Millman, S. J. van der Walt, R. Gommers, P. Virtanen, D. Cournapeau, E. Wieser, J. Taylor, S. Berg, N. J. Smith, R. Kern, M. Picus, S. Hoyer, M. H. van Kerkwijk, M. Brett, A. Haldane, J. F. del Río, M. Wiebe, P. Peterson, P. Gérard-Marchant, K. Sheppard, T. Reddy, W. Weckesser, H. Abbasi, C. Gohlke, and T. E. Oliphant. 2020. Array programming with NumPy. Nature 585:357–362. https://doi.org/10.1038/s41586-020-2649-2.

Kaler, J., G. J. Wassink, and L. E. Green. 2009. The inter- and intra-observer reliability of a locomotion scoring scale for sheep. Vet. J. 180:189–194. https://doi.org/10.1016/j.tvjl.2007.12.028.

Lawin, F. J., A. Byström, C. Roepstorff, M. Rhodin, M. Almlöf, M. Silva, P. H. Andersen, H. Kjellström, and E. Hernlund. 2023. Is markerless more or less? Comparing a smartphone computer vision method for equine lameness assessment to multi-camera motion capture. Animals (Basel) 13:390. https://doi.org/10.3390/ani13030390.

Lidfors, L. 1989. The use of getting up and lying down movements in the evaluation of cattle environments. Vet. Res. Commun. 13:307–324. https://doi.org/10.1007/BF00420838.

Riaboff, L., S. Poggi, A. Madouasse, S. Couvreur, S. Aubin, N. Bédère, E. Goumand, A. Chauvin, and G. Plantier. 2020. Development of a methodological framework for a robust prediction of the main behaviours of dairy cows using a combination of machine learning algorithms on accelerometer data. Comput. Electron. Agric. 169:105179. https://doi.org/10.1016/j.compag.2019.105179.

Riaboff, L., L. Shalloo, A. F. Smeaton, S. Couvreur, A. Madouasse, and M. T. Keane. 2022. Predicting livestock behaviour using accelerometers: A systematic review of processing techniques for ruminant behaviour prediction from raw accelerometer data. Comput. Electron. Agric. 192:106610. https://doi.org/10.1016/j.compag.2021.106610.

Schnitzer, U. 1971. Abliegen, Liegestellungen und Aufstehen beim Rind im Hinblick auf die Entwicklung von Stalleinrichtungen für Milchvieh. Kuratorium Für Technik Und Bauwesen in Der Landwirtschaft-Bauschriften 10:43.

Seabold, S., and J. Perktold. 2010. Statsmodels: Econometric and Statistical Modeling with Python. Pages 92–96 in Proc. 9th Python in Science Conference.

Shepley, E., J. Lensink, and E. Vasseur. 2020. Cow in motion: A review of the impact of housing systems on movement opportunity of dairy cows and implications on locomotor activity. Appl. Anim. Behav. Sci. 230:105026. https://doi.org/10.1016/j.applanim.2020.105026.

Shewbridge Carter, L., S. M. Rutter, D. Ball, J. Gibbons, and M. J. Haskell. 2021. Dairy cow trade-off preference for 2 different lying qualities: Lying surface and lying space. J. Dairy Sci. 104:862–873. https://doi.org/10.3168/jds.2020-18781.

Truong, C., L. Oudre, and N. Vayatis. 2020. Selective review of offline change point detection methods. Signal Processing 167:107299. https://doi.org/10.1016/j.sigpro.2019.107299.

Vasseur, E. 2017. Animal Behavior and Well-Being Symposium: Optimizing outcome measures of welfare in dairy cattle assessment. J. Anim. Sci. 95:1365–1371. https://doi.org/10.2527/jas.2016.0880.

Vázquez Diosdado, J. A., Z. E. Barker, H. R. Hodges, J. R. Amory, D. P. Croft, N. J. Bell, and E. A. Codling. 2018. Space-use patterns highlight behavioural differences linked to lameness, parity, and days in milk in barn-housed dairy cows. PLoS One 13:e0208424. https://doi.org/10.1371/journal.pone.0208424.

Virtanen, P., R. Gommers, T. E. Oliphant, M. Haberland, T. Reddy, D. Cournapeau, E. Burovski, P. Peterson, W. Weckesser, J. Bright, S. J. van der Walt, M. Brett, J. Wilson, K. J. Millman, N. Mayorov, A. R. J. Nelson, E. Jones, R. Kern, E. Larson, C. J. Carey, İ. Polat, Y. Feng, E. W. Moore, J. VanderPlas, D. Laxalde, J. Perktold, R. Cimrman, I. Henriksen, E. A. Quintero, C. R. Harris, A. M. Archibald, A. H. Ribeiro, F. Pedregosa, and P. van Mulbregt., and SciPy 1.0 Contributors. 2020. SciPy 1.0: Fundamental algorithms for scientific computing in Python. Nat. Methods 17:261–272. https://doi.org/10.1038/s41592-019-0686-2.

Zambelis, A., M. Gagnon-Barbin, J. St John, and E. Vasseur. 2019. Development of scoring systems for abnormal rising and lying down by dairy cattle, and their relationship with other welfare outcome measures. Appl. Anim. Behav. Sci. 220:104858. https://doi.org/10.1016/j.applanim.2019.104858.

ORCIDS

Adrien Kroese https://orcid.org/0009-0001-5780-7345
Moudud Alam https://orcid.org/0000-0002-3183-3756
Elin Hernlund https://orcid.org/0000-0002-5769-3958
David Berthet https://orcid.org/0009-0006-3106-0622
Lena-Mari Tamminen https://orcid.org/0000-0001-6781-4533
Nils Fall https://orcid.org/0000-0001-5597-2358
Niclas Högberg https://orcid.org/0000-0002-2672-7924

II

Smart Agricultural Technology 12 (2025) 101205

Contents lists available at ScienceDirect

Smart Agricultural Technology

journal homepage: www.journals.elsevier.com/smart-agricultural-technology

Evaluating the automated measurement of abnormal rising and lying down behaviours in dairy cows using 3D pose estimation

Adrien Kroese [a,*], Niclas Högberg [a], Elena Diaz Vicuna [b], David Berthet [c], Nils Fall [a], Moudud Alam [d], Lena-Mari Tamminen [a]

[a] Department of Clinical Sciences. Faculty of Veterinary Medicine and Animal Science, Swedish University of Agricultural sciences, Ulls väg 26, 756 51 Uppsala, Sweden
[b] Università degli Studi di Torino. Department of veterinary science, Largo Paolo Braccini 2 10095 Grugliasco, Italy
[c] Sony Sweden, Filial to Sony Europe Ltd. (UK), Mobilvägen 221 88 Lund, Sweden
[d] School of Information and Engineering, Dalarna University, Högskolegatan 2, 791 88 Falun, Sweden

ARTICLE INFO

Keywords:
Animal welfare assessment
Free-stall cubicle
3D pose estimation
Rising behaviour
Lying down behaviour
Precision livestock farming

ABSTRACT

The structure of cubicles can hinder cows' movements when transitioning between postures, leading to atypical motion patterns. Assessing posture transitions relies on visual observations. This study presents a framework for complementing these assessments with kinematic measurements using 3D pose estimation. A total 809 rising and 791 lying down posture transitions were recorded over 12 cubicles by 7 synchronized cameras and processed with 3D pose estimation locating the position of the poll, withers, T13 and sacrum. First, the displacement of the keypoints was used to detect phases of the posture transitions. This detection was compared with visual observations of 200 recordings. The average mean absolute difference in detected timestamps between human and machine across all phases was 0.5 s (average $\sigma = 0.7$) and was under 0.9 s for all phases. Second, indicators were scored based on spatial use and duration, and their distribution compared to existing thresholds. We observed that 59.9 % of rising bouts and 29.1 % of lying down bouts exceeded at least one threshold. Rising delay occurred in 2.8 % of rising bouts and backwards crawling in 59.2 %. Lying down duration exceeded the threshold in 28.9 % of bouts, and rear limbs shifting duration in 8.3 %. Side lunge had a binary threshold which was not adapted to continuous sensor data. Finally, we investigated the association between indicators and found distinct dimensions for head lunge and crawling. We conclude that 3D pose is useful to score posture transition indicators, and that several indicators should be used together to capture distinct dimensions.

1. Introduction

Free stall cubicles are designed to encourage cows to lie down rather than stand, and to defecate outside of the bed. Balancing design elements involves a trade-off at the expense of movement opportunities. For instance, neck rails improve hygiene but increase the incidence of abnormal movements [1]. The ability for cows to comfortably transition between postures is an important parameter of cow comfort in stalls [2, 3].

The ability to perform unhindered posture transitions, such as getting up and lying down, is recognized as a critical component of cow welfare and resting [4,5]. Sufficient space and stable footing are needed to perform these transitions smoothly [6]. It has been hypothesised that the ability to comfortably transition between postures promotes the occurrence of lying behaviour [7]. Adequate rest – in terms of duration, frequency and comfort – is important to dairy cows, studies having shown that cows will work to access resting spots [8]. Brouwers et al. [4] found that in cubicles with flexible dividers, which allow for a more ample movements, cows lied down more frequently and that daily lying duration was higher, suggesting that the ability to lie down without obstruction promotes resting behaviour.

Comfortably transitioning between postures extends beyond physical health, these movements are linked with behavioural expressions of comfort and well-being [5]. Cows that struggle with these transitions may experience increased stress and discomfort, which can affect their overall behaviour and productivity. Providing an environment that facilitates posture transitions can lead to increased resting, and to improved welfare outcomes [9,10]. The quality of posture transition movements is used as a welfare assessment indicator, reflecting the comfort offered by the stall [6,11,12].

* Corresponding author.
E-mail address: adrien.kroese@slu.se (A. Kroese).

https://doi.org/10.1016/j.atech.2025.101205
Received 17 November 2024; Received in revised form 14 July 2025; Accepted 14 July 2025
Available online 18 July 2025

In practice, the assessment of posture transition comfort is typically performed visually by a trained assessor, scoring indicators associated with adverse welfare outcomes, such as bumping the head on the cubicle bars [12]. The Welfare Quality assessment framework contains 2 criteria which are the duration of standing-to-lying (STL) posture transitions and collisions with equipment [11]. The Swedish framework Fråga Kon (Växa, Stockholm, Sweden), which is meant as a practical on-farm assessment of welfare through animal-based measures, assesses the quality of lying-to-standing (LTS). Visual evaluation has limitations, mainly low observation frequency, the inability to re-evaluate when scoring live, and the need for the observer to note various behaviours which may occur simultaneously. Observer disagreement does not seem to be a major risk however; for instance Zambelis et al. [12] reported a Kappa of 0.93 at its lowest when assessing abnormal posture transition indicators. The assessment frameworks presented earlier rely on few or single quantitative indicators for each posture transition.

Precision livestock farming (PLF) technology offers an opportunity to monitor posture transition movements continuously, simultaneously and objectively, and to automatically detect abnormalities in posture transitions.

Sensors have already been used to assess posture transitions. Motion capture has been applied to measuring head lunge (the forward displacement of the head) and showed that cows in open packs lunged further when lying down by a mean of 6 cm while using the same total longitudinal space [13]. Motion capture is a gold standard for kinematic measurements of animals [14] but remains impractical in production settings, which may explain the low sample size (n = 5) in the former study [13]. Brouwers et al. [15] developed a machine learning model to detect abnormal lunge movements from accelerometer data. They used annotations by trained observers of the occurrence of abnormal lunges as labels and tri-dimensional acceleration features as input. The accuracy of their model reached up to 74 %, with the class having the highest accuracy being backwards crawling. This metric is encouraging but needs refining for practical implementation. It is important to note that this result is unlikely due to limitations in the model. Rather, the training labels were annotated using ethograms developed for visual observations, in which the same behaviour class can be reflected by vastly different motion patterns [15].

A possible technology to assess kinematic features during posture transitions is pose estimation [16]. A widespread example of applications of pose estimation in detecting bovine kinematic abnormalities is lameness assessment [17,18]. Pose estimation will track the displacement of key anatomical features to quantify indicators of abnormal locomotion [19]. Kinematic assessment with 2D pose estimation, as is commonly done to assess lameness [19–21] relies on straight walks along an assigned path, perpendicular to the camera's line of sight [17]. Such setup with a fixed orientation of the camera is not feasible for assessing posture transitions of several animals in a production setting. The challenge is that the angle between a single camera's field of view and each stall varies with the stall location, distorting joint angles and perspectives. Pose estimation fusion in 3D from multi-view computer vision however is invariant to camera placement [22] and thus offers more flexibility, when sensor placement is constrained by the existing barn design. Importantly for practical application, pose estimation does not rely on markers (unlike motion capture) and applies to all subjects in the scene (all cows in the cubicles being filmed).

From the state-of-the art in visual assessment there are two challenges that sensor-based posture transition assessment could overcome; the difficulty in scoring multiple indicators in a single event and the time needed to assess regularly. We thus propose a method to identify the phases of posture transitions using multi-view fusion of pose estimation in 3D, and detect the occurrences of abnormalities.

The aim of the study was (i) to develop a method to detect successive phases of cows' posture transitions from 3D poses and score comfort indicators during these phases, (ii) to validate the detection against the human eye and assess its robustness to noisy data and (iii) to study the distribution and possible association of posture transition indicators. To do so, we used a Sony multi-camera system (Sony Sweden, Lund, Sweden) to generate 3D poses of dairy cows in a free stall barn during both posture transitions. Using the 3D pose, we detected the different phases of the posture transitions using change-point detection and supervised learning to then compare the detected timestamps to those annotated by human observers. Then, we measured the duration of each phase as well as kinematic features to identify bouts with indicators exceeding thresholds for comfortable movements. Finally, we investigated whether there existed an association between indicators.

2. Materials and methods

In this study, we use 3D pose to measure indicators of posture transition quality. Here is a general overview: video sequences showing posture transition bouts were recorded with synchronized cameras with overlapping fields of view. The multi-camera system was calibrated to determine intersecting lines of sight. The 3D pose of cows was inferred from 2D poses estimated on synchronized frames across several cameras. The displacement of anatomical features of cows was tracked throughout bouts and the timestamp of specific phases was detected and compared with manual annotations. Finally, kinematic indicators of posture transition were measured and compared to existing thresholds.

2.1. Location and animals

2.1.1. Study area

Video recordings from 7 cameras (G3 Bullet, Ubiquiti) were collected on 30 separate days between 2021 and 12–08 and 2022–04–28 at all times of day and night. The cameras were placed around an area of a free-stall barn covering 12 stalls (Cubicle divider cc1800 with rigid head bar, Delaval International, Tumba, Sweden) located next the sorting gate of the automatic milking system (VMS 300, DeLaval International, Tumba, Sweden). The cameras were installed around the rows of stalls, between 2.8 and 3.6 m high, and oriented towards the rows of cubicles so that all cubicles in the study ward, including forward lunge room defined as the 60 cm beyond the head rail, were visible by at least 2 cameras. All recordings were obtained at the Swedish Livestock Research Centre's dairy barn (Uppsala, Sweden).

2.1.2. Animals

The herd comprises Swedish Holstein and Swedish Red cattle housed indoors during the study period but with pasture access between May and September. On average, 51 lactating cows were present simultaneously in the pen, with individuals being added and removed throughout, for a total of 183 different individuals having visited the pen during the study period. The average parity of the animals at the start of data collection was 2 with a mode of 1. Days since calving ranged from 6 to 447 with an average of 149. 7 animals were diagnosed with non-reproductive health disorders during the study. Specifically, 3 cows were treated for mastitis, 1 cow was identified with severe lameness, 1 cow with a hoof inflammation, and 2 cows were diagnosed with paresis. Average individual body condition score as measured by the BCS camera (BCS, DeLaval International, Tumba, Sweden) during the trial was 3.4 ± 0.33 (μ±σ).

Cows are milked robotically with voluntary access up to 12 h until which they are brought to milking if they have not gone voluntarily. Passage through the milking robot's sorting gate is necessary to access feed. Cows underwent claw health inspection and trimming every 6 months

2.2. 3D pose estimation

This study employs a synchronized multi-camera system (Sony Sweden) with known intersecting lines of sight to reconstruct 3D poses from 2D key-point estimates. Each pose comprises the coordinates of anatomical landmarks (head at the poll, highest point of the withers, T13, and sacrum at the uppermost point of the ilium) in an arbitrary coordinate system at a given timestamp. HRNET [23] is used to estimate key-points in 2D for each frame. These poses are then fused to obtain 3D key-points.

Frames are synchronized by reading the frame timestamp in the metadata and using the first frame with a common full second transition as frame 0. Synchronization is maintained throughout the recording of up to 35 s by reading the frame order of arrival in the processing buffer for each camera, recording at the same framerate. The 3D fusion of poses is robust to misalignments of up to 0.5 s for movements corresponding to the velocity of a human walking.

Intrinsic calibration parameters are determined using structure-from-motion algorithm [24]. This step determines the cameras' distortion parameters and ensures alignment of all cameras' origin and axes with world coordinates [25]. Then, the system was extrinsically calibrated to determine intersecting lines of sight between cameras using the technique described by Moliner et al. [26]. A single human is tracked by the pose estimator through the area of interest (twelve cubicles and they alley between them or a surface area of 7.5 × 6.4 m). A preliminary 3D pose of the human is determined by triangulating each unique key-point across 2D poses. The system refines the calibration data through an optimization process that minimizes a reprojection errors function [26]. Reprojection error measures the difference between the observed 2D key-points in the images and the projected 2D locations of the 3D points calculated using the current calibration data. Pose quality assesses the plausibility of the calculated poses based on expected orientations and distances between key-points which have a defined range based on biological constraints (relative position of anatomical key-points to each other). The calibration parameters are then refined iteratively to reduce the reprojection error [26]. The system is robust to temporary occlusions and outliers by using temporal consistency checks.

The system outputs coordinates of the key-points in a 3D space for all objects present in the scene, and associates each keypoint to an object, differentiable by their track number consistent over frames, and a confidence metric (average 2D confidence from HRNET estimation over all 2D poses used to generate the 3D pose). The number of objects is determined by the number of unique key-points. To maintain tracking consistency in assigning key-points to the correct object across timestamps, the system employs a combination of spatial-temporal continuity and trajectory analysis. Once key-points are identified in each frame, the system tracks these points over time by assuming smooth and continuous motion, thereby associating key-points in one frame with their corresponding points in subsequent frames. This process creates trajectories for each keypoint, which are then used to distinguish between different objects based on their unique movement patterns. Additionally, the system incorporates a smooth motion error function during optimization, which penalizes non-uniform acceleration of key-points between frames, further ensuring spatio-temporal consistency. Fig. 1 exemplifies the 3D pose in two separate events by showing the vertical coordinate of key-points during STL for each track.

The pose estimator expresses coordinates in an approximation of the meter. It is important to note that while the scale of units expressed by the 3D key-points is consistent across locations, its exact resolution is unknown. This means that all values given in meters should be considered as $m \pm \mathbb{C}$ where $\mathbb{C}$ is an unknown constant. The implications of this limitation is that great caution should be exercised when comparing absolute values to other studies but that analysis of association and change rates are unaffected.

2.3. *Video sequence selection*

Initially, 979 videos showing a lying-to-standing bout and 1015 showing standing-to-lying were visually identified for development purposes [27] and reused for this study. We applied a simple event detector calculating the difference in average withers Z position (height) across 10 frames (0.3 s) between the start and end of the sequence. An absolute difference above 0.4 m was considered to be a posture transition, and the direction of change (downwards for STL and inversely) informed on the type. This is visible in Fig. 1 where the withers go from a height of about 1.7 m to 1 m.

After detecting events, 814 and 798 sequences were classified as lying to standing and standing to lying respectively. This corresponds to respective false negative rates of 16.9 % and 21.4 %. After visually inspecting the key-point series for each sequence, 5 and 26 sequences were noticed to have been misclassified as LTS and STL and subsequently removed, giving false positive rates of under 1 % and 3.2 %. The sequences contained the 30 to 35 s video recorded by 7 synchronized cameras and show cows transitioning between postures in a cubicle. Removal of false positives left 809 and 791 LTS and STL sequences respectively.

2.4. *Signal processing of 3D pose time series*

2.4.1. *Filtering*

A low-pass filter with a cut-off frequency of 10 Hz was applied to each key-point and its corresponding X, Y and Z coordinates' time series individually. This approach is based on recommendation by Hamäläinen et al. [28] and by Riaboff et al. [29] for noise removal on animal motion data (originally intended for accelerometer data). The filter was implemented in Python 3.9 using the function "butter" from the SciPy package [30].

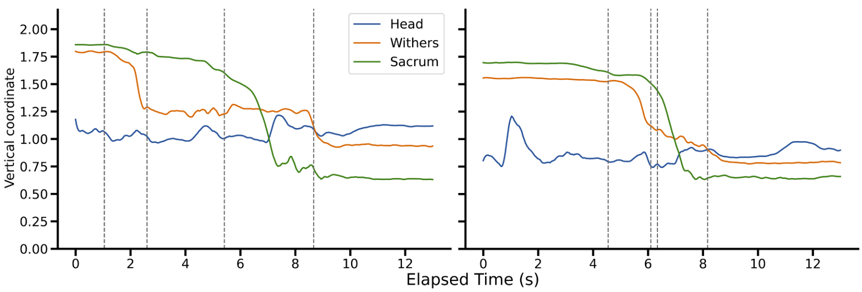


Fig. 1. Vertical coordinate of 3 key-points during two lying down motions, comparing slow with swift posture transitions. Dashed lines correspond to the detection of the initial leg bend, thoracic limbs on ground, sacrum descent and completion. On the right pane, the rapid sacrum descent initiates just before the front limbs touch the ground. These examples were cherry-picked for clarity.

2.4.2. Stitching discontinuous tracks

Object tracking could be interrupted by factors such as noise peaks or temporary occlusions, leading to instances where successive detections of the same animal were split between multiple tracks. To address this issue, we implemented a post-processing track-stitching algorithm that merges fragmented tracks corresponding to the same animal into a single continuous track, based on spatial continuity of the smoothened key-point coordinates. The track-stitching algorithm operates by first identifying all tracks within a given sequence and calculating the time and position at which each track ends. The algorithm then searches for subsequent tracks that begin within a temporal window of 1 s and spatial proximity of 0.3 (in the pose estimator's coordinate system, corresponding approximately to 30 cm). Candidate tracks that start shortly after the end of the previous track are evaluated based on their Euclidean distance in the 3D space, using the wither key-point's coordinate. The algorithm prioritizes merging tracks that are closest in space. Tracks are iteratively processed until no further stitching opportunities are detected. This method resulted in the inclusion of 305 LTS and 301 STL posture transitions sequences, representing 37.7 % and 38.1 % respectively of the total sequences used.

2.4.3. Interpolating missing poses

The tracking algorithm has a tolerance to punctual missing detections and stitched tracks had a gap up to 1 s. This resulted in instances where consecutive 3D poses were separated by more than the expected interval of 0.033 s. To ensure consistency, poses were interpolated for missing frames, thereby standardizing the time intervals between consecutive poses. First, gaps were identified based on the timestamp difference between consecutive poses, and the number of missing frames was calculated. We estimated missing poses using 3D cubic spline interpolation — a method Ren et al. [31] found to be highly faithful for interpolating missing positions in cow movement data—thereby achieving uniform temporal resolution across sequences and facilitating further calculations.

2.5. Indicators of posture transition quality

Indicators relevant to assessing the quality of the posture transition were retrieved from the literature and are listed in Table 1. This study focuses on the movement opportunities offered by the cubicles, and the occurrence of atypical motions. For this reason, inclusion criteria for indicators were (i) measurable during the posture transition movement and (ii) measurable through kinematic features at a specific phase of the posture transition. The start and end of the posture transition movements are described in Table 1. Atypical motions such as dog sitting and horse-like rising were initially selected but did not occur. The selected indicators, their definition and corresponding phase, as well as existing thresholds beyond which the motion is considered abnormal are gathered in Table 1.

Out of the selected indicators, lying down duration, hind quarters shifting, delayed rising, backwards crawling and head lunge space had quantified thresholds found in the literature. Side lunge was described as yes or no in the ethograms found in Brouwers et al. [15] and in Dirksen et al. [32].

2.6. Event detection during posture transition and indicator calculations

To measure the indicators of comfortable posture transition it was necessary to accurately detect the occurrence of specific phases during the motion using the key-points' displacement. These phases are listed in the third column of Table 1.

The main method here is change-point detection in the key-point coordinates, specifically the Y (perpendicular to the stall) and Z (vertical) coordinates of the withers. Change-point detection involves identifying indices in a time series where there is a shift in the series' statistical properties, such as mean or variance. In the case of the key-points 3D coordinates time series, change-points represent movements from one posture to another. The detection process involves segmenting the time series into distinct windows where the statistical properties are consistent within each segment but differ between segments. The change-points are the boundaries of these segments. Linearly penalized segmentation (Pelt) used here [33] optimizes the segmentation by balancing the number of change points against the fit to the data, using a penalty parameter to control the trade-off. The Pelt algorithm is implemented in the Python library Ruptures [34]. Parameters for change-point detection were optimized through a grid search testing the penalties of 3, 5 and 10 with any combination of the x and y coordinates of the withers or sacrum, and their movement velocity. For each combination, the mean absolute difference (MAD) was calculated between the annotated timestamp for that phase and the timestamp

Table 1
Selected indicators of posture transition comfort.

Indicator	Definition	Corresponding phases	Threshold for acceptable comfort
Rising			
Duration of rising motion	Start of the motion: the cow gathers its front limbs under the body causing a visible rise in the withers' position [27] End of the motion: the cow is fully up with all limbs extended [6]	Rising on breastbone, Standing	
Backwards crawling on carpal joints	When resting on carpal joints, the cow moves its front leg backwards before the lunge motion [12]	Rising on breastbone, lunge	None/0 m [12]
Delayed rising	The cow rests on its carpal joints before lunging.	Rising on breastbone, lunge	$<$ 10 s [12]
Head lunge distance	Euclidian distance projected in 2D above the bed, measured between the point of furthest extension of the head and the position of the withers just before the lunge (after possible backwards movements)	Lunge, head baseline location	$>$ 0.6 m beyond the end of the cubicle [22]
Side lunge	Maximum angle formed between the lines joining the poll to the neck and the neck to the t13 during the lunge [27].	Lunge	No side lunge [15,32]
Lying down			
Duration of lying-down motion	Start of the motion: one carpal joint is bent and lowered [11]. End of motion: the cow is fully lying down and the body is stable [12]	Initial leg bend, recumbent position	$<$ 6.3 s [11]
Hind quarter shifting	Duration between the moment both carpal joints touch the ground and the rapid descent of the sacrum.	Thoracic limbs touchdown, sacrum descent	$<$ 3 s [12]
Head displacement	Length of the horizontal vector between the head at start of the movement and its point of furthest forward displacement	Head maximum extension	0.59 m (mean maximum in open pen) [13]

corresponding to the nearest change-point. The variables and penalty creating a change point closest to the annotation are reported in the respective sub-section for each phase.

This method outputs several change-points in each sequence, corresponding to the different phases, as well as other events and also possibly noise. Thus it was necessary to select the right change point corresponding to the phase of interest amongst the various change-points detected.

The velocity of the withers and sacrum display specific patterns in between each phase as the cow moves parts of its body in succession. Thresholds in velocity peaks were used to constrain time windows for each phase and thus select the correct change-point. Rules and thresholds for change-point selection are described in Table 2 and in the subsections dealing with the detection of specific phase. It was not possible to detect all events in all sequences, and the final sample sizes used to calculate each indicator are found as labels on 3 and Fig. 5 respectively in the results section.

Fig. 1 Illustrates two STL sequences on which the timestamps detected for the phases have been marked by dashed vertical lines. On the left panel, the initial drop of the withers (orange curve), corresponding to the leg bend, was detected to have occurred at 9.6 s (first vertical dashed line). This is followed by readjustment movements of the hind quarters while the cow is standing on its thoracic limbs between the 11.8 s and 14.1 s timestamps. This characterised by a plateau of the withers height, as the cow rests on its anterior limbs during the posterior readjustment movements. On the example on the right, the motion is a lot swifter, with only a brief deceleration of the withers' descents, as both anterior limbs reach the ground at 15.2 s.

The methods to detect most phases are listed in Table 2. Other phases as well as kinematic indicators have a dedicated sub-section.

2.6.1. Backwards crawling

Before lunging, when forward space is perceived as insufficient the cow moves its front limbs backwards [12]. Identifying this movement enables to quantify the crawling distance but also enables the establishment of a consistent baseline position of the withers immediately prior to the head lunge, which is crucial for calculating the displacement of the head during the lunge. Backwards crawling was defined as the total backwards displacement of the withers key-point's coordinate along the x axis, between the start of the rising motion and the head lunge.

Table 2
Posture transition phases and methods for detection.

Posture transition phase	Penalty	Variables for change-point detection	Threshold for selecting a cv-hange-point
Rising (LTS)			
Start of rising motion	10	Withers Y, Withers Z	First change point where the median Z withers in the following 1 s window > median Z withers in the initial 1 s of the sequence
Head lunge	Maximum Head X coordinate		
Standing	5	Withers velocity	First change point after the last velocity peak of 0.18 (normalized units)
Lying down (STL)			
Initial leg bend	10	Withers vertical velocity	Last change point before the first peak in withers velocity above 0.2 (normalized units)
Thoracic limbs touchdown	3	Withers Z	First change-point immediately after the first peak above 0.2
Sacrum descent	Random forest		
Recumbent position	10	Withers Y, Withers Z	Last change point where the median Z withers in the following 1 s window < median Z withers in the final 1 s of the sequence

2.6.2. Head displacement and angle

For the analysis of head lunge, sequences were only used if the head key-point maintained a confidence level above 0.77 during lunge. The confidence threshold was decided by plotting the distribution of confidence values of the head around the predicted lunge timestamp, and visually identifying an elbow in the plot.

The withers baseline position was defined as the X coordinate of the withers after backwards crawling, also corresponding to the minimum X coordinate between start of the rising motion and lunge when crawling was not detected to have occurred. Lunge distance was defined as the distance on the x axis between the head at lunge and the withers baseline location, to which was subtracted the distance between the head and the withers at lunge. The rationale behind this calculation was to determine how far forward the head was able to lunge, not compared to the cubicle, but to the initial placement of the cow before lunging.

Head lunge angle was calculated as the 2D projected angle over the horizontal plane, formed by the line of the back (joining the withers to the sacrum) and the neck (joining the withers to head keypoint) at the moment of furthest extension. An angle of 180° represents straight lunge, where the head is exactly aligned with the back. A lower angle represents a sideways neck, independent of lunge side. (Fig. 2)

For the head displacement when lying down, the maximum filtered coordinate of the head on the X axis (parallel to the stalls) was subtracted to the head's position on the X axis at the time of initial leg bend.

2.6.3. Thoracic limb touchdown

This refers to the earliest point at which both anterior limbs are folded and the cow touches the bed with both carpal joints. The withers' coordinate was normalized and their vertical velocity was computed. Change-point detection with a penalty of 3 was applied to the withers Z coordinate series. Peaks in the wither's vertical displacement above 0.2 normalized distance units per second were detected, with a minimum distance between peaks of 40 points or 1.33 s. We selected the first change-point following the peak first.

2.6.7. Sacrum descent

The change-point method failed to produce detections corresponding to the sacrum descent timestamp. Instead, the following methods were tried: recurrent neural network with dropout and one of each 1 dimension convolutional, bi-directional long-short-term-memory and dense layers, against a random forest with 50 estimators predicting the index of the event. The RNN produced a MAE on unseen data between detections and annotations of 0.81 s at the stabilisation of the loss term after 12 epochs while the random forest produced a MAE of 0.41 s and was thus chosen. Since the sequences were of varying length and usually centred on the posture transition, and to avoid overfitting the model to a specific location in the sequence, the key-point series were randomly padded before training the models. Padding was added at the beginning and end of each series, for a total length of 1147 (arbitrary value above the length of the longest series) according to the following equations:

L_{pad_s} is the total padding length for sequence s: $L_{pad_s} = 1100 - L_s$ with L_s being the length of sequence S.

L_{start_s} is the padding length at the start of sequence s: $L_{start_s} \sim Uniform(0,\ L_{pad_s})$

L_{end_s} is the padding length at the end: $L_{end_s} = L_{pad_s} - L_{start_s}$. The padding values are calculated as follows:

$$P_{pos,\ k,\ x} = coord_{pos,k,x} \otimes I_{L_{pos}*1} + N \tag{1}$$

where P is the matrix of padding values of size $6 * L_{pos}$ with pos taking values *start* or *end*, *coord* being the first or last value in the series for coordinate $x = X$ or Z and key-point k = withers or sacrum. $N \sim Uniform(0,\ 0.05)$ is a vector of random noise. Considering S, the original sequence of key-point positions, the padded sequence used as input in

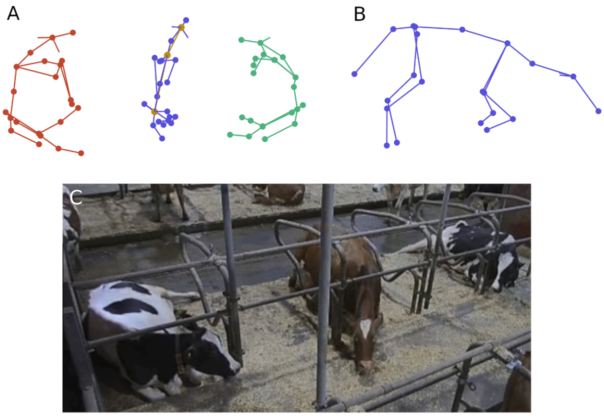


Fig. 2. 3D pose of cows with one cow rising (blue pose) taken at furthest head extension (lunge). Head lunge angle is defined as the angle between the segments in yellow on pane A, joining the head, withers and sacrum (highlighted). A: top down view of all 3D poses in the row of cubicles. B: side view 3D pose of the cow rising. C: corresponding frame.

the random forest is:

$$\begin{bmatrix} P_{start} \\ S \\ P_{end} \end{bmatrix}$$

2.7. Validation

2.7.1. Agreement between observers and with event detection

To validate the accuracy of the detection of the various phases, video sequences showing posture transitions by a single cow were annotated by 3 observers. The observers annotated the timestamps for each event listed in column 3 of Table 1. Observers first trained on 10 sequences for each posture transition and agreed on the timestamps to annotate. Then, each observer was provided with a total of 100 video sequences for each posture transition, which were randomly assigned, shuffled and blinded. The 100 sequences contained 55 videos which were common to all observers. This overlap was to score inter-observer agreement. The 100 videos also contained 30 sequences which were unique to each observer. Among the resulting 85, 15 were randomly resampled to assess intra-observer agreement. For each sequence to be annotated, the material provided to the observers contained the synchronized video from all 7 cameras. Observers were free to choose the camera offering the best view of the cow performing the posture transition.

Agreement was measured as MAD between annotated and detected timestamps. $MAD(i,m) = \frac{1}{300}\sum_{s=1}^{300}\left|\Delta_{s,(o,m)}\right|$ where $\Delta_{s,(o,m)} = \left|t_{s,i} - t_{s,m}\right|$ with m being the automated detection and $t_{s,i}$ the time stamp of the s:th sequence by o:th observer.

2.7.2. Agreement depending on interruptions in the poses

Sequences contained 637 to 1013 consecutive poses, including sequences stitched from spatio-temporally continuous tracks. We ran a regression to analyse the effect of the presence of a stitch in a $\pm$ 1.7 s window around the annotation, as well as the duration of interpolated poses on the agreement between annotations and detections. The model is described as follows:

$$T_{o_e} = \beta_1 + \beta_2 M + \beta_3 S + \beta_4 I + 1|sequence + \varepsilon \quad (2)$$

where T is the observed timestamp, either annotated or detected. e is the event (taking values of all 7 events in both posture transitions). M is the observer type indicating whether the timestamp was annotated by a human or detected by the model. S is a dichotomous variable representing the presence of a track-stitch in the 3D pose sequence. It always takes the value of 0 in the case of human annotation (because stitches in the 3D pose have no meaningful effect on human annotations performed on the video), and 1 or 0 in the case of model detections, depending on the presence of a stitch in the $\pm$1.7 s window around the mean human annotation. The value of 1.7 s corresponds to the 95th percentile of differences between human and machine. Similarly, I is the interpolated duration in case of detections and 0 in the case of annotations. Finally, $1|sequence$ represents a random intercept for the sequence number, as the predicted timestamp in each sequence has no tangible meaning and is relative to the start of the video but should theoretically be equal for all annotations in the same sequence. We report the value and the significance of β_2, β_3 and β_4. β_2 represents the difference in predicted event timestamp if the observation was done by the model compared to a human, β_3 the change in predicted timestamp if the observation was done by the model and a stitch was present in the $\pm$1.7 s window and β_4 the change in predicted timestamp for 1 s of interpolated poses in the window. Significance is accepted at a risk of $\alpha = 0.05$.

2.8. Exclusion criteria

Sequences were first included if a posture transition was detected from the key-point data. This produced 809 sequences in which the occurrence of a LTS posture transition was visually identified, and 791 STL. Regarding annotations, 145 sequences of each posture transition were originally annotated. 4 annotated STL sequences were discarded as well as 4 LTS sequences because of data quality issues.

After the events of interest were detected using the methods described above for the entirety of the sequences, including all sequences which had not been annotated, the validity of the detection was visually assessed using the vertical displacement graphs, of which Fig. 1 shows an example for two different sequences. The time-series of the key points' vertical coordinate were plotted for all sequences, and the detected timestamps were added to the plots. Sequences were excluded based on visual assessment if any of the detected timestamps did not match the kinematic pattern corresponding to the event. 84 LTS and 87 STL sequences were excluded, the number of events that were inaccurately detected is listed in Table 3.

Table 3
Count and frequency of detection errors per event. Note that the total errors amount to more than the total, as several events could be off in the same sequence.

Event	Errors	Frequency
Rising (LTS)		
Rise on breastbone	22	2.7 %
Lunge	34	4.2 %
Standing	37	4.6 %
Any	84	10.4 %
Lying down (STL)		
Initial leg bend	17	2.2 %
Thoracic limbs touchdown	29	3.7 %
Sacrum descent	36	4.6 %
Recumbent position	13	1.6 %
Any	87	11.0 %

2.9. *Statistical analysis of indicator scores*

To explore the association between indicators, Spearman's correlation was calculated between indicators in the same posture transition sequence. A principal component analysis (PCA) was conducted to identify more complex correlations between indicators in LTS transitions using PCA function from SciKit-Learn [35].

3. Results

The purpose of this study was to evaluate the accuracy in the detection of the successive stages in the posture transitions, to detect the occurrence of indicators exceeding thresholds for comfortable posture transition and to explore possible indicator association.

3.1. *Comfort indicators exceeding thresholds*

In the stalls used for this study, and regarding duration, we found that 2.8 % of LTS posture transitions exceeded the threshold for indicator 'Rising delay'. If we use the 5 s threshold used in Fråga Kon, instead of 10 s found by Zambelis et al. [12], 30.2 % of LTS bouts would exceed the threshold. Crawling backwards occurred in 59.2 % of LTS transitions. 28.9 % of STL exceeded the threshold for total duration and 8.3 % for shifting duration. Altogether, 59.9 % of LTS and 29.1 % of STL exceeded thresholds for at least one indicator.

Table 4
Mean absolute difference (in seconds, ± standard deviation) in annotated or detected timestamps for each pair of observers and with the automated detection.

Feature	Observer pair: Obs 1 - 2	Obs 1 - 3	Obs 2 - 3	Observers – machine
Rising (LTS)				
Rise on breastbone	1.1 ± 1.4	1.8 ± 1.6	1.0 ± 1.3	0.9 ± 1.1
Head lunge	0.2 ± 0.3	0.2 ± 0.4	0.2 ± 0.4	0.3 ± 0.6
Standing	0.5 ± 1.1	0.6 ± 1.1	0.4 ± 0.4	0.7 ± 0.9
Lying down (STL)				
Leg bend descent	0.3 ± 0.2	0.2 ± 0.2	0.2 ± 0.2	0.4 ± 0.7
thoracic limbs touchdown	0.2 ± 0.1	0.2 ± 0.1	0.2 ± 0.1	0.4 ± 0.6
Sacrum descent	0.4 ± 0.4	0.4 ± 0.4	0.3 ± 0.4	0.4 ± 0.5
Recumbent position	0.8 ± 0.4	0.5 ± 0.5	0.5 ± 0.6	0.4 ± 0.7

3.2. *Agreement on phase detection and robustness to interrupted poses*

The results in Table 4 show agreement under half a second for most events. The first phase of the rising movement showed the most disagreement between observers. (Table 5)

When missing positions were interpolated, the average interpolated duration was 0.5 *s* ± 0.5 (μ±σ) or 31 % of frames in the window around the event for LTS and 0.7 ± 0.7 or 43 % of frames for STL. For both rising and lying down transitions, interpolating poses on missing frames did not have a significant effect on the timestamp prediction by the model. Only for the rising on breastbone and the thoracic limbs touchdown phases did the presence of a stitch have a significant effect on the difference between annotated and detected timestamps (at $\alpha = 0.05$). The observed timestamp being detected by the model rather than a human observer was only significant for the thoracic limbs touchdown.

3.3. *Distribution and association of posture transition comfort indicators*

The analysis of association between indicators was aimed at understanding whether there existed a combination of indicators which by themselves offer a summary of the posture transition quality, or rather if indicators showed no association and that there was thus no relation between the qualities of the different phases. Both posture transitions were analysed separately.

3.3.1. *Lying to standing*

Rising duration had a median of 8.3s±2.8 (median ± Standard deviation) and a skewness of 1.4. Total duration does not have a threshold on Fig. 3 since no recommendations were found. For rising delay, it was 4.0s±2.4 with a skewness of 1.4. Crawling distance had a median of 0.1 ± 0.1 and a skewness of 1.1. Lunge distance showed an important range from 0.3 to 1.5. Its median was 0.66±0.33 and its skewness 0.44. Lunge angle had a median of 159.7°± 11 and its distribution was skewed to the left (skewness −0.6).

Spearman's pairwise correlations, shown as labels on Fig. 3 revealed a set of moderately to strongly correlated variables ($p < 0.001$): duration, crawling distance and rising delay. Lunge angle and distance had a negligible yet significant correlation ($p = 0.005$), the significance driven by the high sample size ($n = 548$). (Fig. 4, Fig. 5)

The principal component analysis aimed at exploring whether the indicators could be combined into subsets that better explain the movement patterns. The first 4 components were retained, explaining 98 % of the variance in the dataset.

The first component (PC1) explains 45 % of variance. Variables with the highest loading on PC1 were delay, crawling and duration. The second component (PC2) explains 23 % of the variance and is loaded by head lunge distance and angle.

Table 5
Coefficients for the effect of the processing method and event detection on the predicted timestamp based on Eq. (2). Significant coefficients are bolded.

Feature	Coefficient for the effect on predicted timestamp (seconds) (n sequences with processing method): Presence of stitch	Duration of interpolation	Model detection	ICC
Rising (LTS)				
Rise on breastbone	**−1.4** (7)	0.0 (27)	−0.1	0.83
Head lunge	−0.4 (10)	0.1 (72)	−0.0	0.87
Standing	0.3 (14)	0.4 (32)	−0.1	0.78
Lying down (STL)				
Leg bend descent	−0.0 (9)	−0.3 (38)	−0.1	0.91
thoracic limbs touchdown	**−0.5** (18)	0.1 (91)	**−0.4**	0.95
Sacrum descent	−0.1 (29)	0.0 (14)	−0.0	0.94
Recumbent position	0.2 (7)	0.3 (30)	−0.0	0.83

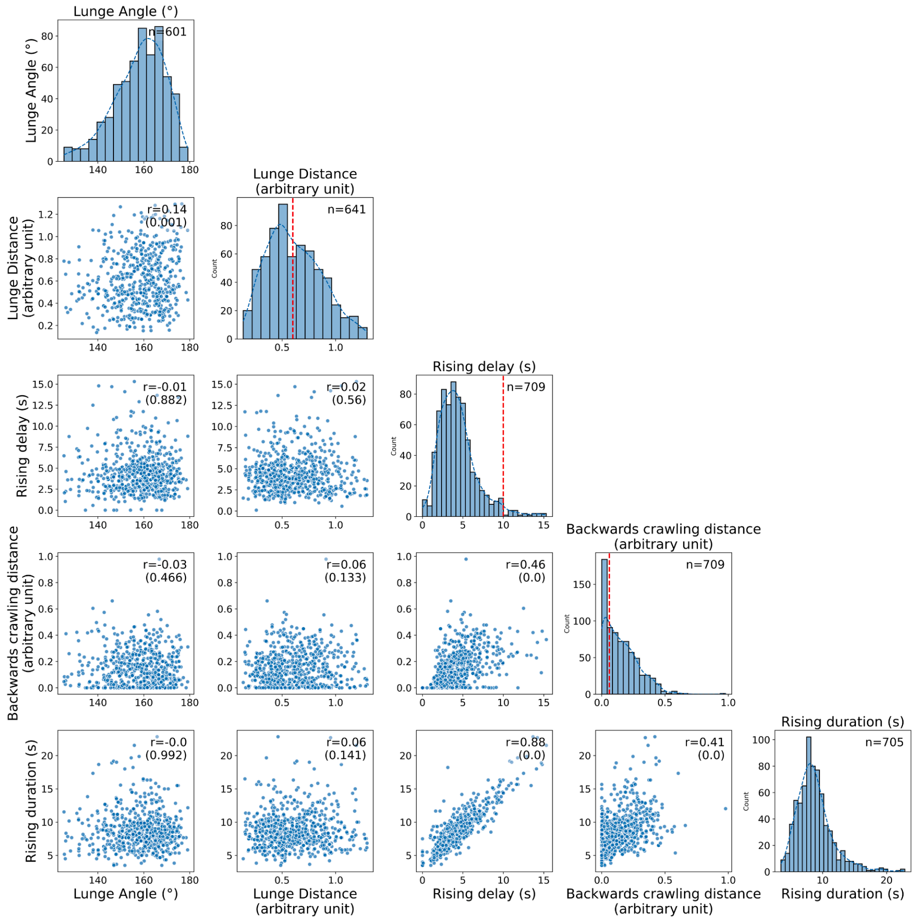


Fig. 3. Distribution with kernel density estimation and pairwise scatterplots of lying to standing posture transition indicators. Cut-offs for comfort assessment are indicated by a dashed red line on the histograms when found in the literature. The sample size for each indicator is reported with the histograms and the correlation (p-value) on the scatterplots.

Components 3 and 4 explained 17 and 13 % of variance respectively. Component 3 had lunge angle and distance load with opposing signs. Component 4 showed opposed loading signs between duration and crawling distance.

3.3.2. Standing to lying

The Spearman correlation between shifting duration and lying down duration is 0.66 ($p < 0.001$). Lying down duration had a median of 5.6 ± 1.7 and a skewness of 2.2 while shifting duration had a median of 1.4 ± 1.2 and skewness of 2.4. The distribution of shifting duration is unbalanced, with a high frequency at 0 because of bouts not displaying a window for hind quarter shifting.

4. Discussion

The study comprises several final and intermediate results, which all have implications for dairy cow comfort monitoring in free stalls using pose estimation in 3D This discussion will first offer a summary of key findings regarding both validation of the method and indicator scores, it will then compare them with earlier research and discuss limitations and implications for future cattle welfare monitoring.

4.1. Validation and agreement with human observation

The results confirmed a high agreement between human and algorithmic detection of posture transition phases. The agreement between human and machine in detecting the timestamp of specific events has two implications. The first one is that we can use the system to measure the duration of the successive phases of the posture transition, which is a comfort indicator. The second implication is that the 3D capture system properly captures the kinematics of events of interest since what is seen on the video matches change points in the 3D coordinates. We note that the development was done with a single cubicle design and that the algorithm may not perform equally well in other systems. Supervised

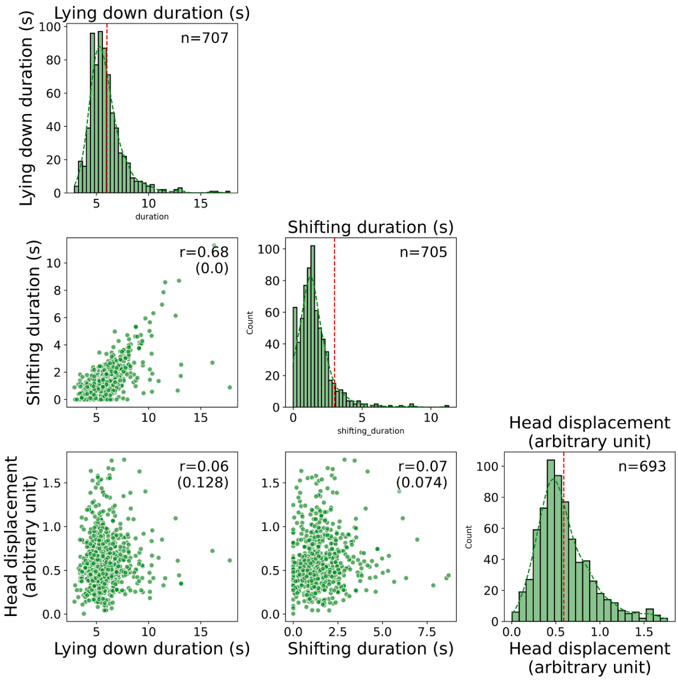


Fig. 4. PCA biplot for lying to standing indicators showing variable loadings and individual scores on 4 components.

learning methods for event detection using diverse sequences can likely address this limitation. The failure rate of up to 11 % does mean that human supervision is required before making meaningful conclusions.

4.2. Comparison of the results with previous studies

Most STL sequences (71 %) were within the accepted duration, however backwards crawling during LTS was highly prevalent. This prevalence comes in stark contrast to results by Brouwers et al. [36] who found a probability of backwards crawling no higher than 5 % in different cubicles. These differences are likely imputable to different stall designs. Zambelis et al. [12] did not find backwards crawling to be associated with characteristics of the cow, nor with adverse welfare outcomes. We still found the indicator be included in the Fråga Kon manual. Combined with the fact that it is rarely observed in unrestricted environments [6], lead us to advocate for its inclusion when evaluating cubicle designs.

Delayed rising; or a pause before the swift head lunge movement, above the suggested threshold had a low prevalence in our study (2.8 %), compared with the 19.5 % reported by Zambelis et al. [12], hinting again at the fact that indicator distributions vary greatly with stall design and thus that the results presented here should not be extrapolated to other farm settings.

The range of forward head displacement shows that even if forward lunge space is offered, cows use this space very differently. We observed on the video that some cows had slow and hesitant movements, with the head not extending beyond the head rail, while others would lunge far forward, potentially explaining the measured variability. Thresholds exist for lunge room, for example 0.9m according to Cook [3]. The 3D coordinates in the system used here however were not precisely expressed in meters. Although the system approximates the meter by design in the calibration phase, caution is warranted when comparing displacement measurements to previous results.

Rising duration is dependent on the identification of the start of the rising motion, which is the phase with the highest ambiguity to observers (over 1 s average difference). Rising duration was positively associated with cow width in the study by Zambelis et al. [12], while delayed rising was not. Delayed rising was a binary indicator in the latter study [12]. Larger cows were predicted to lunge further in an earlier study [37]. A possible explanation for both these results is that larger cows are more hesitant throughout the bout but not specifically before lunge. In the Fråga Kon framework, the threshold for delayed rising was 5 s instead of 10, which would lead to a different observed prevalence.

4.3. Assessing comfort with a combination of indicators and 3D pose

In the Welfare Quality framework, posture transitions are assessed using two indicators; duration and collisions [11]. In their "Flowchart for Evaluating Free Stalls", Nordlund [38] assesses posture transitions through lunge and "bob" spaces, and rising room (measured as the absence of collisions). The manual for Fråga Kon uses the duration of the pause on the front limbs as main indicator. It exemplifies abnormal rising with backwards crawling, dog sitting and difficulty to rise (assessors have also stated looking at side lunge), and gives the expert assessor the discretion to judge, looking at a more complete picture of the cow. Taken separately, indicators provide a simplified view, which is practical for on-farm applications but may not capture the full complexity of the posture transition process.

According to the PCA, there are several uncorrelated patterns of rising motions. PC1 is interpreted as corresponding to hesitation,

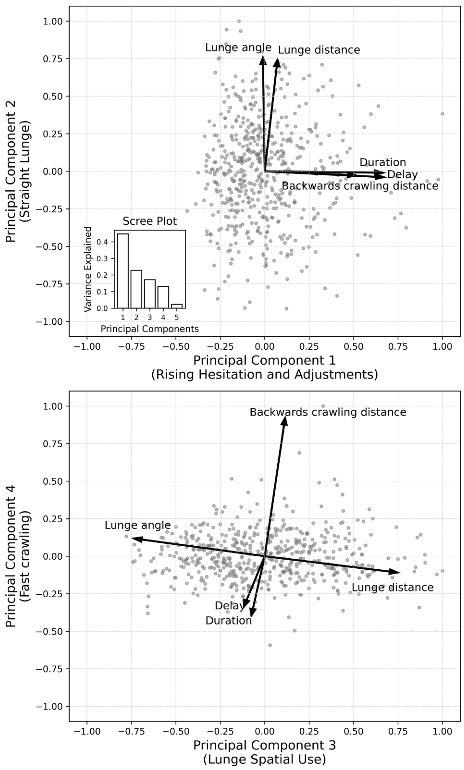


Fig. 5. Distribution with kernel density estimation and scatterplot of standing to lying posture transition indicators. Cut-off for shifting duration is indicated by a dashed red line on the lower histogram. The sample sizes for both indicators are reported with the histograms and the correlation (p-value) on the scatterplot.

creating pauses in the rising motion. This is because the variables loading the highest on PC1 are delay, crawling distance and total duration. These variables are correlated, which is sensible since the further a cow will crawl, the more time it needs to do so, which increases delay and total duration directly. PC2 represents straight lunge, which is a desirable pattern. Lunge distance and angle had a low correlation, but they loaded similarly on PC2, suggesting that they measure distinct but complementary aspects of lunge behaviour. There is seemingly an upper diagonal bound on the scatterplot for these two variables (Fig. 3) which would indicate that angled lunges rarely are associated with longer distance. Components 3 and 4 seem to show exceptions from the most common motion patterns; component 3 had lunge angle and distance load with opposing signs, representing both lateral and longitudinal spatial use while bouts scoring high on PC 4 would represent cows crawling an important distance but quickly.

Principal components being uncorrelated implies that crawling (PC1) is not associated with straight lunge (PC2), contrary to what we had previously hypothesized (the rationale being that crawling backwards offered more forward space to then lunge straight). We know that the stall design in the study farm promotes backwards crawling, which tends to increase delayed rising through readjustments as is reflected in PC1. Loadings on PC4 however show an opposite pattern where cows do crawl but swiftly. Taken together, PC1 and PC4 imply that duration is not systematically an indication of crawling. The first component suggests that the proxy indicator found in Welfare Quality or Fråga Kon are sound summarisation of the parameters explaining the most variability, but the other components suggest that there are additional dimensions to the quality of posture transitions which we should not summarise into a single indicator.

4.4. Defining thresholds based on existing variability and quantitative measurements

The distribution of indicator values presented in the results highlights that the range of posture transition movements, and the duration of the different phases exist on a continuum. This comes in stark contrast with the rigid thresholds found in the literature which may not be adapted to assessment using sensor data, of which lunge angle is a clear example. In a similar development, Brouwers et al. [15] found moderate accuracy (60 %) in detecting the occurrence of side lunge using accelerometer data. While the class for side lunge was yes or no, there seems to exist a continuum of lunge angles as shown on the first density plot of Fig. 3. It is worth exploring if misclassifications happen more consistently when the head lunge is at a slight angle. This would mean that the challenge in classifying side lunge in the latter study is not a shortcoming in the algorithm but rather a limitation in the ethogram used in annotations which is not adapted to continuous data [15]. This might lead to misalignments between the sensor output and the annotation, especially in the range of neck angles that represent the borderline between normal and abnormal lunge angle. Bewley et al. [39] describe side lunge as that performed in cubicles designed specifically to allow for cows to lunge their head side, instead of forward (because the cow could be impeded by a wall or another cow). We saw accordingly, in studies assessing posture transitions, that side lunge was a yes/no indicator [12,15]. In the study presented here, the cubicles were designed for forward lunge. However, we did both observe and record bouts in which the neck was at an angle compared to the head. It is important to define whether this form of angled forward lunge classifies as side lunge, if it is another form of abnormal lunge, or if rather it should not be considered abnormal but an individual preference. Anecdotally expert assessors judged some of the lunges in our study as being sideways, yet we found no apparent cut-off in the distribution of lunge angles. This hints to the fact that side lunge is more complex than forward versus sideways, but that there also may not be a universal threshold for what angle constitutes side lunge. This trend towards not observing clear cut-offs from the distribution is visible in all the indicators measured here. We propose that assessment of posture transitions using sensor data should not be done against a rigid threshold. This technology paired with individual recognition could quantify the variability within the herd and individuals, help understand individual motion patterns and tailor the benchmarks to each cow.

4.5. Limitations and necessary improvements for practical implementation

In our previous study, validating a data processing method to detect the start of the rising motion, using the same key-points, we had excluded sequences for which the rising motion was split into several tracks [27]. In real world settings, data generation mechanisms will inevitably produce gaps. In order to move towards implementing such tools in practice, it was important to test whether interrupted sequences could still provide an accurate detection of the posture transition phases. The results were encouraging and showed that stitching tracks and interpolating poses had little effect on the accuracy of the event detection.

Improvements should be made in the system to obtain coordinates in meters, which would allow comparing lunge room with earlier studies [6,13,40]. This would also help provide recommendations regarding cubicle dimensions based on spatial use [2].

Challenges remain for practical application, namely dealing with inaccuracies in event detection and the high false negative rate. The current detection method was a rule-based approach, which relies on the high interpretability of 3D pose estimation, reflecting the actual movement amplitude and location of the anatomical features. This high interpretability can reduces the amount of annotated data needed for event detection and can be relied upon to verify the validity of the detections by setting numerical constraints based on the assumed relative location of the key-points, to each other and to their previous location. Once we identify the kinematic pattern of a phase, we can split longer key-point time series into windows and find matching patterns.

For the head lunge space threshold, we used an average forward displacement of 0.6 m reported by [6]. It is consistent with the findings of [13] who reported a mean maximum displacement of 0.59 m when lying down. This however remains an average and quantifying the variability within the herd is instrumental in designing stall elements which can accommodate all cows.

More posture transition indicators exist than were used in this study. A detailed list can be found in Zambelis et al., [12]. This study was, limited to kinematic indicators.

5. Conclusion

This study showed that 3D fusion of pose estimation is a possible sensor technology to complement posture transition assessment with kinematic measurements. It shows good accuracy on detecting events, with disagreements with human-made visual observations being under 0.5 s for most phases and 0.9 s at most. Human oversight is needed for final evaluation since up to 11 % of sequences had at least one incorrect detection.

Measuring posture transition indicators showed that over half of rising events and under a third of lying down events were considered abnormal. Backwards crawling before rising was particularly prevalent in the farm and cubicles studied.

Analysing the association of indicators with a PCA showed that the dimensions of lunge, hesitation and spatial use were uncorrelated. Backwards crawling, delay, and head lunge should be assessed through specific indicators to cover these distinct dimensions separately. In practice, this is challenging to perform visually, and pose estimation offers a method to increase the information available to assessors.

Notes

The study presented here was approved by the ethical committee Uppsala djurförsöksetiska nämnd under approval 5.8.18–13,069/2021.

The authors thank the Swedish Research Council (Formas) for funding this research (grant 2021–02,254), the personnel of the Swedish Dairy Research Center at Lövsta for their help, and Sony Sweden for the extensive collaboration.

Sony Sweden provided the technology to generate the 3D pose. Conceptualization, study design, statistical analysis and presentation of the results were decided by researchers at the Swedish University of Agricultural Sciences. Sony Sweden contributed in drafting the methods section regarding key-point acquisition in 3 dimensions, and in revising the final manuscript.

Generative artificial intelligence (AI) was used to rephrase paragraphs and improve readability as well as to generate code for the analyses.

Ethical statement

This manuscript involves non-invasive research on 183 animals (*Bos taurus*). The authors declare that all procedures in the submitted work were conducted in accordance with national regulations regarding the use of purpose-bred research animals. The research was approved by the ethical committee Uppsala djurförsöksetiska nämnd under approval 5.8.18-13069/2021.

CRediT authorship contribution statement

Adrien Kroese: Writing – review & editing, Writing – original draft, Visualization, Validation, Methodology, Investigation, Formal analysis, Conceptualization. **Niclas Högberg:** Writing – review & editing, Validation, Supervision, Resources, Project administration, Methodology, Investigation, Funding acquisition, Conceptualization. **Elena Diaz Vicuna:** Writing – review & editing, Methodology, Investigation. **David Berthet:** Software, Resources, Data curation, Conceptualization. **Nils Fall:** Writing – review & editing, Supervision, Project administration, Funding acquisition. **Moudud Alam:** Writing – review & editing, Visualization, Validation, Supervision, Methodology, Formal analysis. **Lena-Mari Tamminen:** Writing – review & editing, Visualization, Supervision, Methodology, Funding acquisition, Formal analysis.

Declaration of competing interest

The authors declare the following financial interests/personal relationships which may be considered as potential competing interests:

Sony provided the technology to generate the 3D pose. Conceptualization, study design, statistical analysis and presentation of the results were decided by researchers at the Swedish University of Agricultural Sciences. Sony contributed to drafting the methods section regarding key-point acquisition in 3 dimensions, and in revising the final manuscript. This study was not performed with the intent of supporting a commercial or competitive claim.

Data availability

The authors do not have permission to share data.

References

[1] J. St. John, J. Rushen, S. Adam, E. Vasseur, Making tiestalls more comfortable: I. Adjusting tie-rail height and forward position to improve dairy cows' ability to rise and lie down, J. Dairy Sci. 104 (3) (2021) 3304–3315, https://doi.org/10.3168/JDS.2019-17665.

[2] N.B. Cook, Free-stall design for maximum cow comfort, WCDS Adv. Dairy Technol. 21 (21) (2009) 255–268. https://acrre.ualberta.ca/wp-content/uploads/sites/57/wcds_archive/Archive/2009/Manuscripts/FreeStallDesign.pdf.

[3] N.B. Cook, Optimizing resting behavior in lactating dairy cows through freestall design, Vet. Clin. North Am.: Food Animal Practice 35 (1) (2019) 93–109, https://doi.org/10.1016/J.CVFA.2018.10.005.

[4] S.P. Brouwers, M. Simmler, M.F. Scriba, P. Savary, Cubicle design and dairy cow rising and lying down behaviours in free-stalls with insufficient lunge space, Animal (2024) 101314, https://doi.org/10.1016/J.ANIMAL.2024.101314.

[5] S.S. Nielsen, J. Alvarez, D.J. Bicout, P. Calistri, E. Canali, J.A. Drewe, B. Garin-Bastuji, J.L.G. Rojas, C.G. Schmidt, M. Herskin, V. Michel, M.Á.M. Chueca, B. Padalino, H.C. Roberts, H. Spoolder, K. Stahl, A. Velarde, A. Viltrop, A.D.B. des Roches, C. Winckler, Welfare of dairy cows, EFSA J. 21 (5) (2023), https://doi.org/10.2903/j.efsa.2023.7993.

[6] L. Lidfors, The use of getting up and lying down movements in the evaluation of dairy cattle environments, Vet. Res. Commun. 13 (4) (1989) 307–324, https://doi.org/10.1007/BF00420838.

[7] C.B. Tucker, D.M. Weary, Stall design: enhancing cow comfort, Adv. Dairy Technol. 13 (2011) 155–167. https://www.researchgate.net/publication/255658372.

[8] M.B. Jensen, L.J. Pedersen, L. Munksgaard, The effect of reward duration on demand functions for rest in dairy heifers and lying requirements as measured by demand functions, Appl. Anim. Behav. Sci. 90 (3–4) (2005) 207–217, https://doi.org/10.1016/J.APPLANIM.2004.08.006.

[9] D.B. Haley, J. Rushen, A.M. de Passillé, Behavioural indicators of cow comfort: activity and resting behaviour of dairy cows in two types of housing, Can. J. Anim. Sci. 80 (2) (2000) 257–263, https://doi.org/10.4141/A99-084.

[10] C. Tucker, D. Weary, D. Fraser, Free-stall dimensions: effects on preference and stall usage, J. Dairy Sci. 87 (5) (2004) 1208–1216, https://doi.org/10.3168/jds.S0022-0302(04)73271-3.

[11] H. Blokhuis, M. Miele, I. Veissier, B. Jones, in: H. Blokhuis, M. Miele, I. Veissier, B. Jones (Eds.), Improving Farm Animal Welfare, Brill | Wageningen Academic, 2013, https://doi.org/10.3920/978-90-8686-770-7.

[12] A. Zambelis, M. Gagnon-Barbin, J.S. John, E. Vasseur, Development of scoring systems for abnormal rising and lying down by dairy cattle, and their relationship

with other welfare outcome measures, Appl. Anim. Behav. Sci. 220 (2019), https://doi.org/10.1016/j.applanim.2019.104858.

[13] A. Ceballos, D. Sanderson, J. Rushen, D.M. Weary, Improving stall design: use of 3-D kinematics to measure space use by dairy cows when lying down, J. Dairy Sci. 87 (7) (2004) 2042–2050, https://doi.org/10.3168/JDS.S0022-0302(04)70022-3.

[14] F.J. Lawin, A. Byström, C. Roepstorff, M. Rhodin, M. Almlöf, M. Silva, P. H. Andersen, H. Kjellström, E. Hernlund, Is markerless more or less? Comparing a smartphone computer vision method for equine lameness assessment to multi-camera motion capture, Animals 13 (3) (2023) 390, https://doi.org/10.3390/ani13030390.

[15] S.P. Brouwers, M. Simmler, P. Savary, M.F. Scriba, Towards a novel method for detecting atypical lying down and standing up behaviors in dairy cows using accelerometers and machine learning, Smart Agric. Technol. 4 (2023) 100199, https://doi.org/10.1016/J.ATECH.2023.100199.

[16] C.-H. Lee, T. Mendoza, C.-H. Huang, T.-L. Sun, Vision-based postural balance assessment of sit-to-stand transitions performed by younger and older adults, Gait Posture 117 (2025) 245–253, https://doi.org/10.1016/j.gaitpost.2025.01.001.

[17] X. Kang, X.D. Zhang, G. Liu, A review: development of computer vision-based lameness detection for dairy cows and discussion of the practical applications, Sensors 21 (3) (2021) 753, https://doi.org/10.3390/s21030753.

[18] A. Nejati, A. Bradtmueller, E. Shepley, E. Vasseur, Technology applications in bovine gait analysis: a scoping review, PLoS ONE 18 (1) (2023), https://doi.org/10.1371/journal.pone.0266287.

[19] K. Zhao, M. Zhang, J. Ji, R. Zhang, J.M. Bewley, Automatic lameness scoring of dairy cows based on the analysis of head- and back-hoof linkage features using machine learning methods, Biosys. Eng. 230 (2023) 424–441, https://doi.org/10.1016/J.BIOSYSTEMSENG.2023.05.003.

[20] A. Poursaberi, C. Bahr, A. Pluk, A.V. Nuffel, D. Berckmans, Real-time automatic lameness detection based on back posture extraction in dairy cattle: shape analysis of cow with image processing techniques, Comput. Electron. Agric. 74 (1) (2010) 110–119, https://doi.org/10.1016/J.COMPAG.2010.07.004.

[21] K. Zhao, J.M. Bewley, D. He, X. Jin, Automatic lameness detection in dairy cattle based on leg swing analysis with an image processing technique, Comput. Electron. Agric. 148 (2018) 226–236, https://doi.org/10.1016/J.COMPAG.2018.03.014.

[22] Ma, H., Chen, L., Kong, D., Wang, Z., Liu, X., Tang, H., Yan, X., Xie, Y., Lin, S.-Y., & Xie, X. (2021). TransFusion: cross-view fusion with transformer for 3D human pose estimation. *ArXiv*. https://doi.org/10.48550/arXiv.2110.09554.

[23] Wang, J., Sun, K., Cheng, T., Jiang, B., Deng, C., Zhao, Y., Liu, D., Mu, Y., Tan, M., Wang, X., Liu, W., & Xiao, B. (2019). *Deep high-resolution representation learning for visual recognition* (Version 2). arXiv. https://doi.org/10.48550/ARXIV.1908.07919.

[24] S.J. Maybank, O.D. Faugeras, A theory of self-calibration of a moving camera, Int. J. Comput. Vis. 8 (2) (1992) 123–151, https://doi.org/10.1007/bf00127171.

[25] E.R. Davies, The dramatically changing face of computer vision. Advanced Methods and Deep Learning in Computer Vision, Elsevier, 2022, pp. 1–91, https://doi.org/10.1016/b978-0-12-822109-9.00010-2. ISBN 9780128221099.

[26] O. Moliner, S. Huang, K. Astrom, Better prior knowledge improves human-pose-based extrinsic camera calibration, in: 2020 25th International conference on pattern recognition (ICPR), 2021, pp. 4758–4765, https://doi.org/10.1109/ICPR48806.2021.9411927.

[27] A. Kroese, M. Alam, E. Hernlund, D. Berthet, L.-M. Tamminen, N. Fall, N. Högberg, 3D pose estimation to detect posture transition in free-stall housed dairy cows, J. Dairy Sci. 107 (9) (2024), https://doi.org/10.3168/jds.2023-24427.

[28] W. Hamäläinen, M. Järvinen, P. Martiskainen, J. Mononen, November). Jerk-based feature extraction for robust activity recognition from acceleration data, in: 11th International Conference on Intelligent Systems Design and Applications, 2011, https://doi.org/10.1109/ISDA.2011.6121760.

[29] L. Riaboff, S. Poggi, A. Madouasse, S. Couvreur, S. Aubin, N. Bédère, E. Goumand, A. Chauvin, G. Plantier, Development of a methodological framework for a robust prediction of the main behaviours of dairy cows using a combination of machine learning algorithms on accelerometer data, Comput. Electron. Agric. 169 (2020) 105179, https://doi.org/10.1016/J.COMPAG.2019.105179.

[30] P. Virtanen, R. Gommers, T.E. Oliphant, M. Haberland, T. Reddy, D. Cournapeau, E. Burovski, P. Peterson, W. Weckesser, J. Bright, S.J. van der Walt, M. Brett, J. Wilson, K.J. Millman, N. Mayorov, A.R.J. Nelson, E. Jones, R. Kern, E. Larson, Y. Vázquez-Baeza, SciPy 1.0: fundamental algorithms for scientific computing in Python, Nat. Methods 17 (3) (2020) 261–272, https://doi.org/10.1038/s41592-019-0686-2.

[31] K. Ren, M. Alam, P.P. Nielsen, M. Gussmann, L. Rönnegård, Interpolation methods to improve data quality of indoor positioning data for dairy cattle, Front. Animal Sci. 3 (2022), https://doi.org/10.3389/fanim.2022.896666.

[32] N. Dirksen, L. Gygax, I. Traulsen, B. Wechsler, J.-B. Burla, Body size in relation to cubicle dimensions affects lying behavior and joint lesions in dairy cows, J. Dairy Sci. 103 (10) (2020) 9407–9417, https://doi.org/10.3168/jds.2019-16464.

[33] R. Killick, P. Fearnhead, I.A. Eckley, Optimal detection of changepoints with a linear computational cost, J. Am. Stat. Assoc. 107 (500) (2012) 1590–1598, https://doi.org/10.1080/01621459.2012.737745.

[34] C. Truong, L. Oudre, N. Vayatis, Selective review of offline change point detection methods, Signal Proces. 167 (2020) 107299, https://doi.org/10.1016/J.SIGPRO.2019.107299.

[35] L. Buitinck, G. Louppe, M. Blondel, F. Pedregosa, A. Mueller, O. Grisel, V. Niculae, P. Prettenhofer, A. Gramfort, J. Grobler, R. Layton, J. Vanderplas, A. Joly, B. Holt, G. Varoquaux, API design for machine learning software: experiences from the scikit-learn project, arXiv. (2013), https://doi.org/10.48550/ARXIV.1309.0238.

[36] S.P. Brouwers, A.F.E. Schug, M. Simmler, P. Savary, The effect of neck strap positioning relative to dairy cow body size on rising, lying down, and defecation behaviour in lying cubicles, Animal (2025) 101507, https://doi.org/10.1016/j.animal.2025.101507.

[37] A. Kroese, N. Högberg, D. Berthet, L.-M. Tamminen, N. Fall, M. Alam, Exploring the link between cow size and sideways lunging using 3D pose estimation, in: 11th European Conference on Precision Livestock Farming, 2024, pp. 32–39. https://www.diva-portal.org/smash/record.jsf?pid=diva2%3A1916993&dswid=7796.

[38] K. Nordlund, A flowchart for evaluating dairy cow free stalls, Bovine Practitioner 37 (2) (2003), 89–87.

[39] J.M. Bewley, L.M. Robertson, E.A. Eckelkamp, A 100-year review: lactating dairy cattle housing management, J. Dairy Sci. 100 (12) (2017) 10418–10431, https://doi.org/10.3168/JDS.2017-13251.

[40] U. Schnitzer, Abliegen, liegestellungen und aufstehen beim rind im hinblick auf die entwicklung von stalleinrichtungen für Milchvieh, Kuratorium Für Technik Und Bauwesen in Der Landwirtschaft-Bauschriften 10 (1) (1971) 43.

Acta Universitatis Agriculturae Sueciae

Doctoral Thesis No. 2026:6

Pose estimation in 3D was used to track dairy cows' movements when rising and lying down in cubicles. I explored biomechanical measures, including weight distribution modelling, bringing insights into new welfare indicators that can hardly be approached with visual observations. Applied experimentally, 3D poses revealed that replacing rigid metal rails with flexible straps improved the involvement of the head. There is still great variability between animals, but the work proposes improved cubicles that level the field for welfare of the herd.

Adrien Kroese, holds an Engineer's degree from ISARA-Lyon, France, and an MSc in Organic Agriculture from Wageningen University, the Netherlands.

Acta Universitatis Agriculturae Sueciae presents doctoral theses from the Swedish University of Agricultural Sciences (SLU).

SLU generates knowledge for the sustainable use of biological natural resources. Research, education, extension, as well as environmental monitoring and assessment are used to achieve this goal.

ISSN 1652-6880
ISBN (print version) 978-91-8124-203-4
ISBN (electronic version) 978-91-8124-223-2