J. Dairy Sci. 98:1–12
http://dx.doi.org/10.3168/jds.2015-9428

Development and validation of a visual body condition scoring system for dairy goats with picture-based training

A. Vieira,*[1] S. Brandão,†‡ A. Monteiro,§ I. Ajuda,* and G. Stilwell*
*Animal Behaviour and Welfare Research Laboratory, Centre for Interdisciplinary Research in Animal Health, Faculdade de Medicina Veterinária, Universidade de Lisboa, Avenida da Universidade Técnica, 1300-477 Lisboa, Portugal
†Electrical and Computers Engineering, Instituto Superior Técnico, Universidade de Lisboa, Avenida Rovisco Pais, 1049-001 Lisboa, Portugal
‡Electrical and Computers Engineering, Carnegie Mellon University, 5000 Forbes Ave., Pittsburgh, PA 15213
§Escola Superior Agrária de Viseu, Instituto Politécnico de Viseu, Quinta da Alagoa, Estrada de Nelas, 3500-606 Viseu, Portugal

ABSTRACT

Body condition scoring (BCS) is the most widely used method to assess changes in body fat reserves, which reflects its high potential to be included in on-farm welfare assessment protocols. Currently used scoring systems in dairy goats require animal restraint for body palpation. In this study, the Animal Welfare Indicators project (AWIN) proposes to overcome this constraint by developing a scoring system based only on visual assessment. The AWIN visual body condition scoring system highlights representative animals from 3 categories: very thin, normal, and very fat, and was built from data sets with photographs of animals scored by a commonly used 6-point scoring system that requires palpation in 2 anatomical regions. Development of the AWIN scoring system required 3 steps: (1) identification and validation of a body region of interest; (2) sketching the region from photographs; and (3) creation of training material. The scoring system's reliability was statistically confirmed. An initial study identified features in the rump region from which we could compute a set of body measurements (i.e., measures based on anatomical references of the rump region) that showed a strong correlation with the assigned BCS. To validate the result, we collected a final data set from 171 goats. To account for variability in animal size and camera position, we mapped a subset of features to a standard template and aligned all the rump images before computing the body measurements. Scientific illustrations were created from the aligned images of animals identified as representative of each category to increase clarity and reproducibility. For training material, we created sketches representing the threshold between consecutive categories. Finally, we conducted 2 field reliability studies. In the first test, no training was given to 4 observers, whereas in the second, training using the threshold images was delivered to the same observers. In the first experiment, interobserver results was substantial, showing that the visual scoring system is clear and unambiguous. Moreover, results improved after training, reaching almost perfect agreement for the very fat category. The visual body condition scoring system is not only a practical tool for BCS in dairy goats but also shows potential to be fully automated, which would enhance its use in welfare assessment schemes and farm management.

Key words: body condition scoring, dairy goat, visual body condition scoring system, representative image, template alignment

INTRODUCTION

Goats mobilize body fat reserves according to their nutritional status, physiological needs, and availability of adipose tissue (Morand-Fehr, 2005). Fat mobilization leads to metabolic changes that may have a substantial effect on the animal's health, welfare, and production (Caroprese et al., 2009; Roche et al., 2009). Different methods exist to monitor changes in body fat reserves but BCS is the most widespread method across species, being considered simple and repeatable (Ferguson et al., 2006).

For cattle, sheep and goats, the scoring systems most commonly used to assess BCS are numerical rating scales with 5-point (Edmonson et al., 1989; Ferguson et al., 1994), 6-point (Russel et al., 1969; Lowman et al., 1976; Hervieu and Morand-Fehr, 1999; Harwood, 2006), or 8-point (Wildman et al., 1982) scales. These scoring systems are usually divided into intermediate scores (0.25 or 0.5) that result in 13- to 21-point scoring systems. For this reason, observers need considerable training and experience. Moreover, all of these methods are subjective as they rely on the observer's assessment, which is generally based on visual appraisal or palpation of anatomic locations that may demand

Received February 5, 2015.
Accepted May 28, 2015.
[1]Corresponding author: ana.lopesvieira@gmail.com

animal restraint, which is time consuming and not easily achieved on large farms (Edmonson et al., 1989; Halachmi et al., 2008).

In goats, the most widespread method is that first published in a preliminary form (Santucci and Maestrini, 1985) and later fully presented by Hervieu and Morand-Fehr (1999). It is a 6-point scoring system with intermediate scores (0.25) that requires palpation of 2 anatomical regions: the sternum and the lumbar vertebrae. The validity and interobserver reliability of this scoring system were confirmed by Santucci et al. (1991).

Body condition score is considered a valid, reliable, and feasible welfare indicator with high potential to be included in on-farm welfare protocols (Winckler et al., 2003; Botreau et al., 2009). One of the main challenges in welfare assessment is to develop scientifically sound on-farm protocols focusing on animal-based indicators (such as BCS) that can be implemented in an acceptable amount of time (Knierim and Winckler, 2009).

The inclusion of traditional BCS systems in on-farm welfare protocol for goats presents 2 challenges: (1) very detailed scoring systems require very experienced assessors; and (2) body palpation requires restraint of goats.

The first challenge has been addressed in other on-farm protocols in goats (Anzuino et al., 2010), as well as in other species (Welfare Quality, 2009a,b), by limiting the scoring systems to 2 or 3 levels; for example, very thin, normal, and very fat. Because our main objective was to ensure feasibility when using BCS as a welfare indicator, this simplification was also selected as our priority.

However, the second challenge remained unaddressed. Generally, BCS levels are qualitative scoring systems built with verbal descriptions of how observers should assess visually and palpate different body regions and anatomical features. Most of the time, these verbal descriptors are difficult to elaborate and even more difficult to understand. Recently, different studies in dairy cows and buffalo species have shown that BCS can be obtained from images (Ferguson et al., 2006; Bewley et al., 2008; Negretti et al., 2008). We believe that visual scoring systems that provide a realistic contextual representation of the different levels should be validated as a reliable and replicable method for BCS.

Thus, visual scoring systems should convey the appearance of a representative animal from each category. The definition of a representative for a given category to classify subsequent sets of new images is a common research topic in computer vision. Approaches focus either on the statistical distribution of features (Tu and Yuille, 2004); for example, identifying mean/median images, or on the identification of threshold features; for example, images that are on the transition between 2 different categories (Burges, 1998). We believe that assessors, after becoming acquainted with the spectrum of possible levels and thresholds, can easily identify the category of a new animal by comparing it with the representative category.

Thus, this study had 3 objectives. First, to provide a systematic approach to the development of a visual body condition scoring system for adult dairy goats; second, to develop a training program using the new scoring system based on the concept of threshold images; and finally, to assess the interobserver reliability under field conditions.

MATERIALS AND METHODS

The general flow for the development of a visual body condition scoring system is presented in Figure 1. The study was organized in 2 stages. The first was the preliminary scoring system development study, and it took place between November and December 2012. The second stage, the final study for scoring system development, occurred between February 2013 and July 2014. Data were analyzed using the base packages of the R statistical language (R Core Team, 2013). All the image-processing tasks were performed in Matlab (The MathWorks Inc., 2012).

Preliminary Study for Scoring System Development

Data Acquisition. We used 32 goats [20 Saanen, 5 Alpine, and 7 crossbreed (Saanen × Alpine)] from an intensive dairy goat farm to select the body region that would give us the best representation of BCS levels. The goats were housed in pens with straw litter and had occasional access to exterior exercise areas. Diet was composed of a TMR distributed twice a day, and goats were milked twice a day in an automatic milking parlor (average of 2.0 L per goat per day). Kids were separated from their mothers after birth.

Two experienced assessors independently scored the 32 goats using the BCS method developed by Hervieu and Morand-Fehr (1999). If there was any discrepancy, the final score was decided by consensus. Only adult goats were scored, independently of milk yield and days in milk. The BCS median was 3, with a range of 1.50 to 5.00.

Using a camera (Nikon D60; Nikon, Tokyo, Japan) with automatic exposure adjustment and focus, we photographed the front (sternum region) and the back (rump region) of each goat, as these are the anatomical regions where palpation is performed for BCS evaluation (Hervieu and Morand-Fehr, 1999). Environmental conditions (e.g., light and background) and distance