Body Condition Score Data of Holstein Dairy Cattle: Reproducibility and Trend across Measurements

Agus Susanto[1*], Datta Dewi Purwantini[1], Setya Agus Santosa[1], Dewi Puspita Candrasari[1]

[1]*Faculty of Animal Science, University of Jenderal Soedirman, Purwokerto Indonesia*
[*]*Corresponding author. Email: agus.susanto0508@unsoed.ac.id*

ABSTRACT

Body condition score (BCS) data is generally scored by more than rater and the data is recorded several times after calving. The objective of the study was to quantify the reproducibility measurements among raters and its trend across time of measurements. The body condition score data of Holstein Cows were extracted from the database provided by the Indonesia National Breeding Centre for Dairy and Forage. The BCS data were recorded on January-December 2015 and January-April 2016, which was 4,709 data altogether rated by 4 raters. An R package of psych was used for analyzing the raters reproducibility in which intra-class correlation coefficient (ICC) of type 3 (a model with average fixed raters) was used. The statistical analysis was performed with R program. The mean BCS of rater 1, 2, 3 and 4 were 3.12, 3.08, 3.13 and 3.12, respectively being the average BCS across raters of 3.11. The minimum and maximum of BCS were 2.5 and 4.0, respectively. As much as 14.36 % of cows were categorized to be below standard BCS. The BCS trend across months when the cows were recorded shows a slight declining pattern, however the determination coefficients of the linear regression model were very low (0.02). The intraclass correlation among raters were 0.62-0.64 and 0.87-0.88 when BCS means was included in the computation. The determination coefficient of regression lines of mean BCS on each rater were 0.62-0.80. The study concludes that the reproducibility measurements among BCS data rated by the available raters were moderate to high. The reliability of BCS data among raters could be increased for example by providing intensive training to the raters. The BCS trend over month of measurements was relatively stable. The contribution of each rater to the mean of BCS across raters variation was not equal.

Keywords: *Dairy cows, Classifiers, Consistency, BBPTUHPT Baturraden*

1. INTRODUCTION

Body condition score is routinely and individually recorded in most of dairy farm since this trait plays important role in deducing the status of energy balance of the animals. It is used by breeders as the crucial management aid in order to achieve the optimal milk production and reproduction efficiency. The occurrence of metabolic as well as peri-partum abnormalities can be reduced. Even though BCS is scored in different scale [1] it is internationally agreed that BCS or change of BCS across different time of measurement is beneficial for adding management decision regarding to feeding. It has been reported in the literatures that BCS has association with different traits such as health [2], milk production [3] and live weight [4]. The association between calving and lactation BCS has been reported to associate with ketosis [2], milk fever [1] and displaced abomasum [5].

Different scoring system of BCS have been introduced differently such as 1-5 scale [6], 1-9 scale [7] and 1-10 scale [1] which all agree that the lowest score is being emaciated and the highest score is being obese. Though, [8] recommended that whatever scale of BCS used, the mean value of BCS should be close to (maxscore-minscore)/2. In addition, according to [8] the BCS standard deviation should be about (maxscore-minscore+1) / 6. Maxscore and minscore are the highest and lowest score of BCS observed on the population, respectively. Extreme low (or high) of BCS score is related with health risk [9] and reduced reproductive efficiency [10].

BCS score is a subjective trait where the value was dependent on the rater. To reduce bias, BCS data generally were scored by more than one observers/raters and the individuals cow BCS are obtained from the average of all participating observers/raters. Raters consistency of BCS scoring can be measured in the form of intraclass correlation coefficients [11]. Reliability index or reproducibility of BCS data scored by a few particular raters are done usually by intraclass correlation. It is computed from an analysis of variance by taking the ratio of between-raters variance and its total variance. Ideal body condition score from calving till the animal is dried (**Figure 4**) has been proposed by [12] showing that the BCS declines up to the six weeks of post calving and then increase gradually until the cow is dried. Thus the objectives of the current study were to quantify the magnitude reproducibility of BCS data scored by the 4 available raters in the Indonesia national breeding centre of dairy cattle and forage (BBPTUHPT Baturraden) and its trend across time of measurements.

2. MATERIALS AND METHODS

Data of individual BCS were collected from the national breeding centre for dairy cattle and forage (*Balai Besar Pembibitan Ternak Unggul dan Hijauan Pakan Ternak*, BBPTUHPT) Baturraden, Indonesia. The BCS data analysed in the study were recorded on January-December 2015 and January-April 2016 on Holstein cows which gave birth of at least once. Altogether, there were 4709 individuals BCS data rated by 4 raters included in the analysis. Individuals BCS data were recorded using 1-5 scale based on scoring on pins, thurl, hooks, sacral ligament and short ribs [6]. Scoring of BCS on dairy cows is illustrated on **Figure 1**.

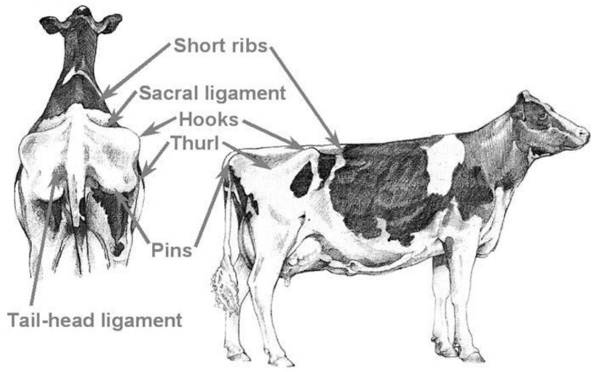

Figure 1. Anatomical areas for body condition scoring (https://www.infovets.com/healthycowinfo/A084.htm)

Reproducibility of BCS data was determined using intra-class correlation coefficient (ICC) of type 3 (a model with average fixed raters). Intra-class correlation coefficient (ICC) was computed using psych package [15]. The BCS trend across month of measurements was fitted using lm package. The graphical illustration of some aspects of the study was created using tidyverse package [16]. All statistical analyses were performed using R program [17].

3. RESULTS AND DISCUSSION

The BCS across raters has the mean ±_standard deviation of 3.11 ± 0.17. The distribution of body condition score of each rater is illustrated in **Figure 2**. It is clear that the BCS data scored by all raters are relatively skewed, though the skewedness direction was inconsistent. The dispersion of the BCS data was identical. The basic statistical values were presented on *Table 1*. The similar dispersion between BCS data scored by different raters could be due to the similar scoring *skill between raters*. This could be due to the rater training for scoring the BCS or due to the working experience of the raters. The staffs assigned as the BCS rater usually have more than 10 years of working experience. The similar dispersion could also be caused by the rule applied to the raters that the score obtained by each rater could not be more than 0.25 compared to those of other raters. If it is the case, the rater must re-score the cow until the requirement for score agreement achieved.

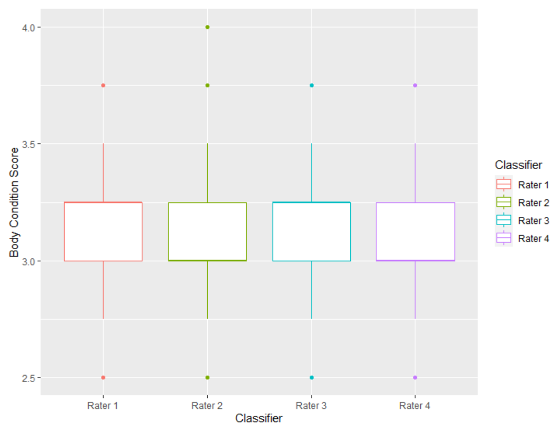

Figure 2. Distribution of body condition score for each rater

The reliability or reproducibility of BCS data among raters using different models of the current study is presented on Table 2. It shows that the intra-class correlation of different models ranges 0.62-0.64; the intra-class correlation is 0.87-0.88 if the means of the available raters is included. In the case of BCS rating where the number of raters involved are fixed as applied to the BCS data of current study, the reproducibility data are categorized moderate (062 - 0.64) depending on the models used in the computation. However, the raters reproducibility is categorized high (0.87-088) if means of the available raters is taken care in the computation.

Table 1. Basic statistics of BCS data

Classifier	Minimum	Maximum	Mean	Median	Standard deviation
Rater 1	2.50	3.75	3.12	3.25	0.16
Rater 2	2.50	4.90	3.08	3.00	0.16
Rater 3	2.50	3.75	3.13	3.25	0.18
Rater 4	2.50	3.75	3.12	3.00	0.16

Table 2. Intra-class correlations computed from pysch package

Models	type	ICC	F	df1	df2	p	lower bound
Single_raters_absolute	ICC1	0.62	7.6	4708	14127	0	0.61
Single_random_raters	ICC2	0.62	8.1	4708	14124	0	0.60
Single_fixed_raters	ICC3	0.64	8.1	4708	14124	0	0.63
Average_raters_absolute	ICC1k	0.87	7.6	4708	14127	0	0.86
Average_random_raters	ICC2k	0.87	8.1	4708	14124	0	0.86
Average_fixed_raters	ICC3k	0.88	8.1	4708	14124	0	0.87

Note: output from pysch package of R program

Cohen's kappa (κ) and spearman's rho (ρ) [18] show relatively similar result of inter-observer agreement. Pearson correlation [19] show that the correlations between raters/scorers are 0.85-0.87 and 0.62-0.66 when the inter-observer reliability is based on the weighted kappa coefficients.

The moderate (or high) intra-class correlation between raters of the current study could be due to the training experienced by the raters. However, the restriction applied to the raters is only to report the allowable value causing this moderate intra-class correlation. As the BCS data are subjective, the restriction applied during scoring could be a disadvantage since the referenced score is not guaranteed to be the most accurate score. To increase the reproducibility of BCS score among raters, multiple observers/raters should be applied [18].

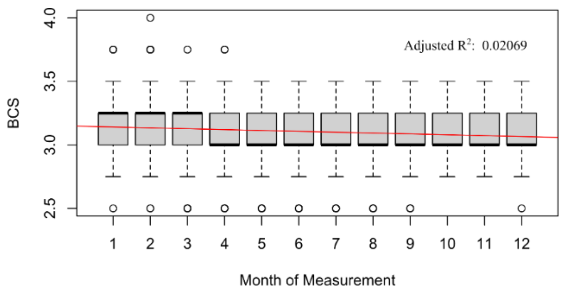

Figure 3. Regression of BCS on month of measurement

The trend of BCS over month of measurementis slightly negative as shown in **Figure 3** being the adjusted R^2 very small (0.02069). Under the dynamic population, the trend of BCS over month of measurement is suggested to be flat since the reproductive activity of the cows is not limited. The small value of R^2indicate that the negative trend of BCS over month of measurement cannot be used to infere the real BCS of the population. The BCS mean within month depends on the structure of the reproductive stage of the cows as it has been pointed out by [12] that normal BCS will follow what is called the ideal function. The deviation of individual BCS could be addressed by either under or over feeding. The negative energy balance and decreased of dry matter intake for instance could increase BCS before and at calving [20] and hence strategies to increase dry matter intake as well as maintaining proper energy balance are important in dairy management.

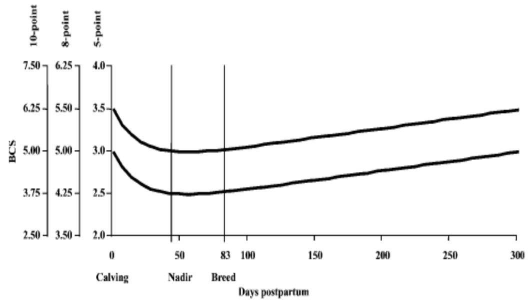

Figure 4. Ideal proposed body condition score from post calving until dry period

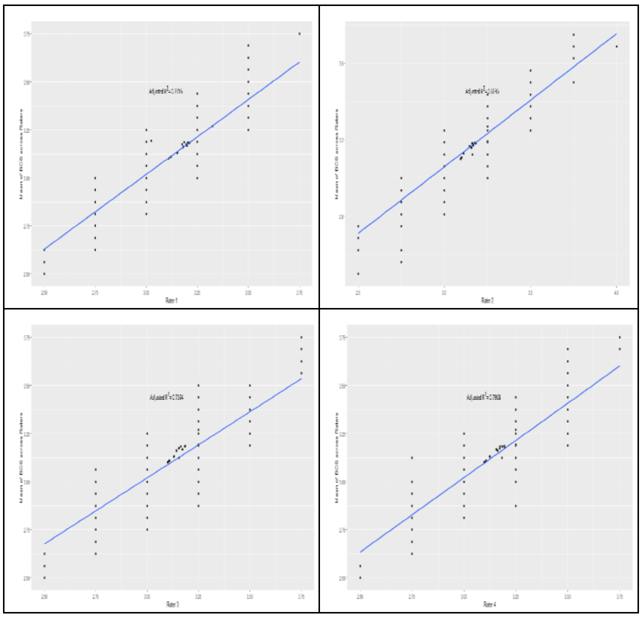

Figure 5. Regression line of across raters BCS mean on each rater

Figure 5 shows the different contribution of each rater to the variation of mean BCS across raters. In general, the variation of mean BCS across raters accounted for by rater is similar (0.74-0.80) except for rater 2 (0.62). It could indicate that the BCS score recorded by rater 2 is less similar than those of the others. This result might be used to evaluate what caused this discrepancy. The BCS rating experience and the training background that rater 2 has joined could be the cause so that providing more on those aspects to rater 2 might increase the scoring accuracy. The subjectivity of raters in rating/scoring the BCS could be avoided by introducing amore advance technology such as digital imaging. The potential of utilizing digital images in assessing body condition score in dairy cows has been studied [21, 22] and the automated BCS rating/scoring might result in more objective and less time-consuming means [23] in obtaining the energy status in dairy cows. Furthermore, [19] have reported a high correlation (0.78 and 0.76) between the automated camera scoring and the manual scoring conducted by three experienced staffs on 343 cows using both continuous and categorical cameras, respectively.

4. CONCLUSION

The reproducibility of BCS data in the National Breeding Centre for Dairy Cattle and Forage of Indonesia (BBPTUHPT Baturraden) was moderate to high. The trend of BCS data over the month of measurement was relatively stable. The contribution of each rater to the mean of BCS across raters variation was not equal. The reliability of BCS data among raters could be increased for example by providing intensive training to the raters.

ACKNOWLEDGMENTS

The authors acknowledge with gratitude the National Breeding Centre for Dairy Cattle and Forage of Indonesia (BBPTUHPT Baturraden) for permitting authors to access the database for this study.

REFERENCES

[1] E.M. Clarke, E.A. Emerson, Design and synthesis of synchronization skeletons using branching time temporal logic, in: D. Kozen (Eds.), Workshop on Logics of Programs, Lecture Notes in Computer Science, vol. 131, Springer, Berlin, Heidelberg, 1981, pp. 52–71. DOI: https://doi.org/10.1007/BFb0025774

[2] J.P. Queille, J. Sifakis, Specification and verification of concurrent systems in CESAR, in: M. Dezani-Ciancaglini and U. Montanari (Eds.), Proceedings of the 5th International Symposium on Programming, Lecture Notes in Computer Science, vol. 137, Springer, Berlin, Heidelberg, 1982, pp. 337–351. DOI: https://doi.org/10.1007/3-540-11494-7_22

[3] C. Baier, J-P. Katoen, Principles of Model Checking, MIT Press, 2008.

[4] M. Kwiatkowska, G. Norman, D. Parker, Stochastic model checking, in: M. Bernardo, J. Hillston (Eds.), Proceedings of the Formal Methods for the Design of Computer, Communication and Software Systems: Performance Evaluation (SFM), Springer, Berlin, Heidelberg, 2007, pp. 220–270. DOI: https://doi.org/10.1007/978-3-540-72522-0_6

[5] V. Forejt, M. Kwiatkowska, G. Norman, D. Parker, Automated verification techniques for probabilistic systems, in: M. Bernardo, V. Issarny (Eds.), Proceedings of the Formal Methods for Eternal Networked Software Systems (SFM), Springer, Berlin, Heidelberg, 2011, pp. 53–113. DOI: https://doi.org/10.1007/978-3-642-21455-4_3

[6] G.D. Penna, B. Intrigila, I. Melatti, E. Tronci, M.V. Zilli, Bounded probabilistic model checking with the muralpha verifier, in: A.J. Hu, A.K. Martin (Eds.), Proceedings of the Formal Methods in Computer-Aided Design, Springer, Berlin, Heidelberg, 2004, pp. 214–229. DOI: https://doi.org/10.1007/978-3-540-30494-4_16

[7] E. Clarke, O. Grumberg, S. Jha, et al., Counterexample-guided abstraction refinement, in: E.A. Emerson, A.P. Sistla (Eds.), Computer Aided Verification, Springer, Berlin, Heidelberg, 2000, pp. 154–169. DOI: https://doi.org/10.1007/10722167_15

[8] H. Barringer, R. Kuiper, A. Pnueli, Now you may compose temporal logic specifications, in: Proceedings of the Sixteenth Annual ACM Symposium on the Theory of Computing (STOC), ACM, 1984, pp. 51–63. DOI: https://doi.org/10.1145/800057.808665

[9] A. Pnueli, In transition from global to modular temporal reasoning about programs, in: K.R. Apt (Ed.), Logics and Models of Concurrent Systems, Springer, Berlin, Heidelberg, 1984, pp. 123–144. DOI: https://doi.org/10.1007/978-3-642-82453-1_5

[10] B. Meyer, Applying "Design by Contract", Computer 25(10) (1992) 40–51. DOI: https://doi.org/10.1109/2.161279

[11] S. Bensalem, M. Bogza, A. Legay, T.H. Nguyen, J. Sifakis, R. Yan, Incremental component-based construction and verification using invariants, in: Proceedings of the Conference on Formal Methods in Computer Aided Design (FMCAD), IEEE Press, Piscataway, NJ, 2010, pp. 257–256.

[12] H. Barringer, C.S. Pasareanu, D. Giannakopolou, Proof rules for automated compositional verification through learning, in Proc. of the 2nd International Workshop on Specification and Verification of Component Based Systems, 2003.

[13] M.G. Bobaru, C.S. Pasareanu, D. Giannakopoulou, Automated assume-guarantee reasoning by abstraction refinement, in: A. Gupta, S. Malik (Eds.), Proceedings of the Computer Aided Verification, Springer, Berlin, Heidelberg, 2008, pp. 135–148. DOI: https://doi.org/10.1007/978-3-540-70545-1_14