IOP Conference Series: Earth and Environmental Science

PAPER • **OPEN ACCESS**

Optimizing Ongole Grade Cattle weight prediction through Principal Component Analysis

To cite this article: A Dakhlan *et al* 2024 *IOP Conf. Ser.: Earth Environ. Sci.* **1360** 012027

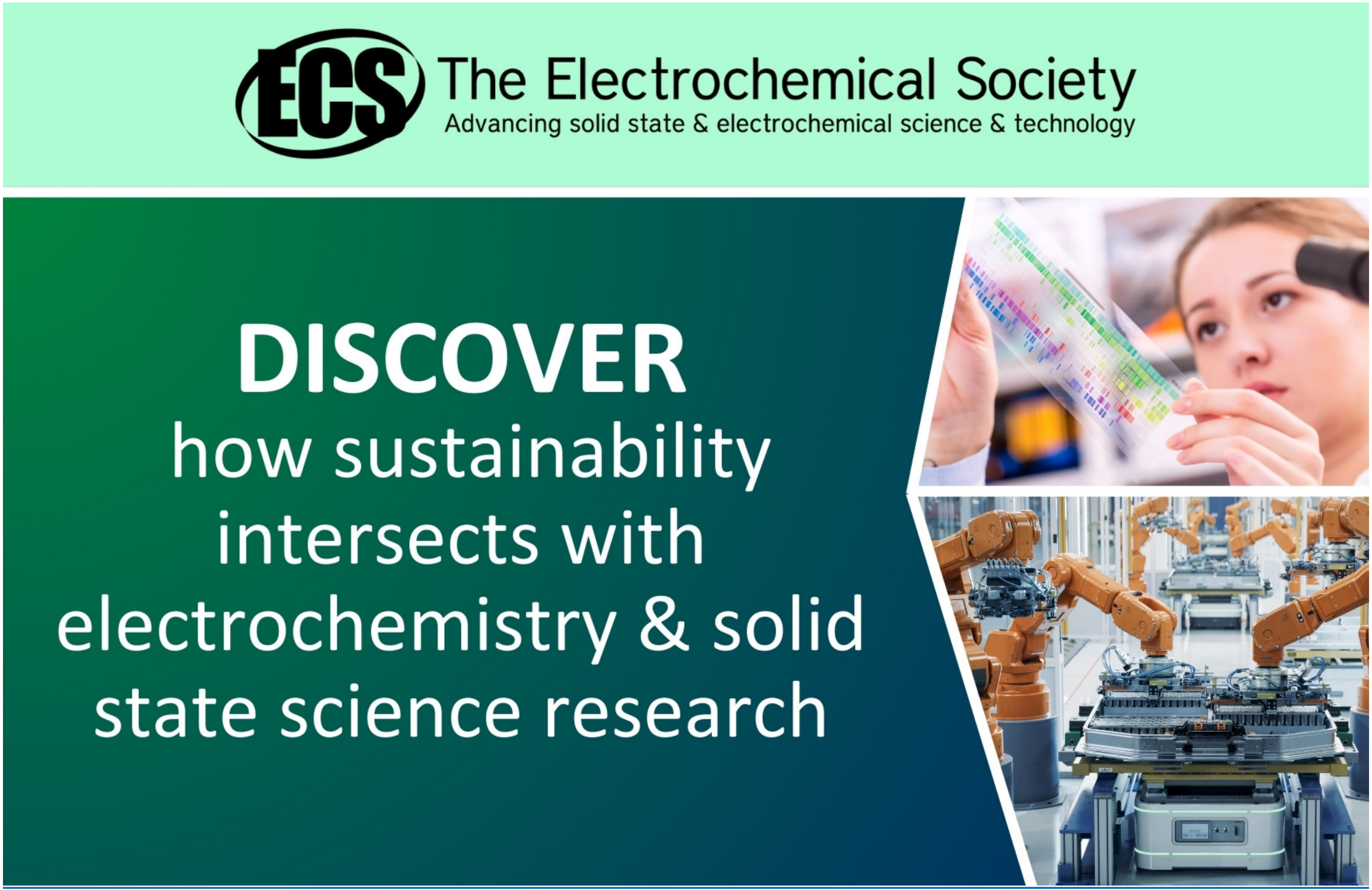
View the article online for updates and enhancements.

You may also like

- Electrochemistry of Secondary Amine Substituted 2,5-di(2-thienyl)pyrrole Derivative and Its Copolymer
Ogun Gumusay, Tugba Soganci, Hakan Can Soyleyici et al.

- Optogalvanic spectroscopy of the hyperfine structure of weak La I lines: discovery of new even parity fine structure levels
Imran Siddiqui, Shamim Khan, B Gamper et al.

- Electrochemical Behavior and LC-MS Analysis of Anthocyanin's in Vaccinium Arctostaphylos L. Extract: The Molecular Modelling of Potential Inhibition to COVID-19 and ROS Generation Receptors
Mahdi Jamshidi, Sara Torabi, Mansoureh Tavan et al.



This content was downloaded from IP address 110.137.36.194 on 19/06/2024 at 15:55

Optimizing Ongole Grade Cattle weight prediction through Principal Component Analysis

A Dakhlan[1], K Adhianto[1], A Qisthon[1], and A Husni[1]

[1]Department of Animal Husbandry, Faculty of Agriculture, Universitas Lampung, Bandar Lampung, Lampung, Indonesia

Corresponding author: akhmad.dakhlan@fp.unila.ac.id
ORCID ID: 0000-0002-8438-6777

Abstract. Ongole Grade (OG) cattle are commonly raised by many farmers in Indonesia, particularly in rural and remote areas where weighing scales are not readily available. The weight of these OG cattle can be estimated by utilizing their body measurements through the application of Principal Component Analysis (PCA). This research aimed to predict the body weight (BW) of OG cattle, which are primarily kept by smallholder farmers, using PCA based on various body measurements such as body length (BL), chest girth (CG), shoulder height (SH), and chest width (CW). Additionally, the effectiveness of PCA-based predictions was compared with a multiple linear regression model. This study involved a total of 120 OG cattle, comprising 26 males and 94 females. The PCA of the body measurements, as well as the correlation and regression between these measurements (BL, CG, SH, and CW) and BW, were analyzed using the R programming language. The selection criteria for identifying the best-fit model for BW prediction were based on statistical indicators including coefficient of determination (R^2), Adjusted R^2, residual standard error (RSE), Akaike Information Criterion (AIC), and Bayesian Information Criterion (BIC). The findings of this investigation revealed that the primary factors representing body measurements were PC_1 for both male (accounting for 89.21% variance) and female OG cattle (accounting for 85.71% variance). In comparison to the regression models, those generated from three PCs for males and two PCs for females demonstrated greater precision and simplicity in estimating the BW of OG cattle, without encountering multicollinearity issues. Consequently, the outcomes of this study have practical applications, serving as a means to predict the BW of OG cattle and contributing to selection programs within this cattle breed.

1. Introduction

Accurate estimation of body weight (BW) is of paramount importance in the management and husbandry of Ongole Grade (OG) cattle, a breed widely reared in Indonesia, particularly in rural and remote areas where resources such as weighing scales are often scarce. The ability to predict the weight of these cattle based on their body measurements is not only a practical necessity but also a crucial factor in enhancing breeding and selection programs, as well as overall livestock productivity.

Traditionally, predicting cattle weight has relied on linear regression models that use individual body measurements like body length (BL), chest girth (CG), shoulder height (SH), and chest width (CW) as predictors. For example, many studies on estimating livestock body weight using linear regression

Content from this work may be used under the terms of the Creative Commons Attribution 4.0 licence. Any further distribution of this work must maintain attribution to the author(s) and the title of the work, journal citation and DOI.
Published under licence by IOP Publishing Ltd

models have been done such as in goats [1-3], sheep [4-6], and cattle [7-12]. Dakhlan *et al* [1] reported combination of chest girth (CG) and body length (BL) was the fittest predictors for body weight (BW) of Ettawa Grade goat with regression model of BW = -67.86 + 0.87*CG + 0.51*BL with coefficient of determination (R^2) 0.76 and for body weight of female Saburai goat with regression model of BW = -36.09 + 0.31*CG + 0.72*BL with R^2 = 0.941. Paputungan et al. [9] suggested that the body weight of OG cattle can be estimated using body measurements (combination of CG and BL) of the cattle with the regression equation of BW = -806.41 + 4.79835*CG + 2.83500*BL with a high degree of accuracy (R^2=0.97). However, these approaches can be limited by the potential for multicollinearity among these variables and may not fully capture the complex relationships that exist within the data, and this will cause redundancy in estimating the body weight of the livestock and this could cause the estimated regression coefficient of the model not reliable and its variance will be high.

Principal Component Analysis (PCA) offers an innovative and promising alternative for optimizing the prediction of OG cattle weight. By transforming correlated body measurements into uncorrelated principal components (PCs), PCA allows for a more efficient and informative representation of the data. This method offers the potential to increase the precision of weight predictions while reducing model complexity and the risk of multicollinearity.

Research on utilizing principal component analysis of body measurements to forecast livestock body weight was still quite limited, particularly in the context of Indonesia. According to Negash's findings [13], the implementation of principal component analysis significantly improved the accuracy of body weight prediction in indigenous Ethiopian chickens. In a similar vein, Canaza-Cayo *et al* [14] also noted that when predicting the body weight of Corriedale Ewes in Southern Peru, a regression model derived from PCA outperformed a regression model that directly used body measurements.

Hence, the objective of the present study was to determine the most effective method for predicting the body weight of Ongole Grade cattle by employing principal component analysis of body measurements and comparing its results with those of the conventional multiple linear regression approach.

2. Materials and Method

2.1. Data collection

Data of body weight and body measurements, including chest girth (CG), body length (BL), shoulder height (SH), and chest width (CW) were obtained from 120 adults OG cattle (26 male and 94 female cattle) aged 3-5.5 years in livestock farmer group of KPT Maju Sejahtera in Tanjung Sari district Lampung Selatan regency, Lampung province, Indonesia. Body weight of cattle was from weighing directly using weighing scale, CG was measured by looping meter tape on cattle chest just behind foreleg, BL was measured as the distance between the shoulder bone bump and sitting bone bump. Shoulder height was measured as the distance between ground and the highest shoulder of the cattle, while CW was measured as distance between left side and right side of chest just behind the foreleg.

Multiple linear regression analysis and PCA were performed using R program [15]. The best-fit prediction of a model was determined by the highest of coefficient of determination (R^2) and adjusted R^2, but the lowest of residual standard error (RSE), Akaike Information Criterion (AIC), and Bayesian Information Criterion (BIC). In addition, stepwise regression was also applied to find the best fit and more parsimonious model. Multiple linear regression of BW from body measurements was formulated as follows.

$$BW = b_0 + b_1CG + b_2BL + b_3SH + b_4CW$$

where b_0 was the intercept, $b_{1\text{-}4}$ was partial regression coefficient for CG, BL, SH, and CW, respectively. While multiple linear regression generated from PCA was formulated as follows.

$$BW = b_0 + b_1PC_1 + b_2PC_2 + b_3PC_3 + b_4PC_4$$

where b_0 was the intercept, $b_{1\text{-}4}$ was partial regression coefficient for PC_1, PC_2, PC_3, and PC_4 scores, respectively.

3. Results and Discussion

3.1. Body measurements and body weight of Ongole Grade cattle

The overview (in boxplot) of body measurements and body weight of Ongole Grade cattle can be seen in Figure 1. Based on the boxplot we can see that body weight, body length, and chest width are normally distributed, while chest girth and shoulder height had outliers. However, based on normality test, all data were normally distributed.

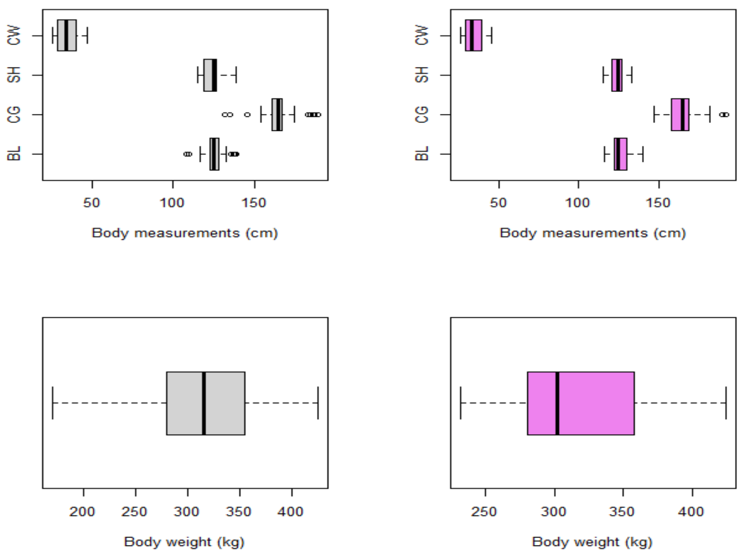


Figure 1. Boxplot of body measurements (BL = body length, CG = chest girth, SH = shoulder height, and CW = chest width) and body weight (BW) of male (left side, grey colour) and female OG cattle (right side, violet colour)

Table 1. Pearson's correlations between body weight and body measurements of male OG cattle (above diagonal) and female OG cattle (below diagonal)

	BL (cm)	CG (cm)	SH (cm)	CW (cm)	BW (kg)
BL (cm)	1.00	0.95**	0.90**	0.84**	0.94**
CG (cm)	0.82**	1.00	0.86**	0.82**	0.96**
SH (cm)	0.80**	0.81**	1.00	0.77**	0.85**
CW (cm)	0.85**	0.83**	0.75**	1.00	0.91**
BW (kg)	0.88**	0.89**	0.82**	0.92**	1.00

Notes: **Correlation is significant (P<0.01); BL = body length, CG = chest girth, SH = shoulder height, CW = chest width, BW = body weight

Table 1 presented the Pearson's correlations among body measurements and BW. Based on Table 1 it can be seen that relationship among body measurements and body weight were strong ranged between 0.75 (CW-SH) and 0.96 (CG-BW) and significant (P<0.01) for both male and female OG cattle. Similar observations were reported by previous study that correlation among body measurement and body weight in Bali cattle were strong and significant [8]. Paputungan *et al* [9] reported that body measurements (CG) highly and significantly correlated to body weight in OG cattle, but among body measurements (BL-CG) was low positive correlations.

3.2. Principal component analysis (PCA)

The principal component (PC) scores of body measurements of male and female OG cattle with their cumulative proportion, proportion of variance, and eigen values is presented in Table 2, and individual PCs is shown in Figure 2 and Figure 3 for male and female OG cattle, respectively. Based on Table 2, Figure 2, and Figure 3 it can be seen that the first PCs accounted for about 89.21 % and 85.71% for male and female cattle, respectively. Furthermore, the eigen value for the first PCs was more than one for the two sex of cattle. This result suggests that the regression models using only one PCs both in male and female OG cattle are adequate for predicting their BW. However, regression models using two or three PCs both in male and female OG cattle would increase the accuracy in predicting BW of OG cattle. The two PCs (PC_1 and PC_2) in this study accounted for about 95.31% and 92.11% of the variance in male and female OG cattle, respectively, and the first three PCs (PC_1, PC_2, and PC_3) accounted for 98.77% and 96.98% for male and female cattle, respectively.

Table 2. The Principal components scores, cumulative proportion, proportion of total variance, and eigenvalues for male and female OG cattle

Male OG cattle	PC_1	PC_2	PC_3	PC_4
Standard deviation	1.889	0.494	0.37211	0.22206
Proportion of variance	0.8921	0.061	0.03462	0.01233
Cumulative proportion	0.8921	0.9531	0.98767	1
Eigen value	3.568228	0.243998	0.138466	0.049309
Female OG cattle	PC_1	PC_2	PC_3	PC_4
Standard deviation	1.8516	0.50573	0.42791	0.36428
Proportion of variance	0.8571	0.06394	0.04578	0.03318
Cumulative proportion	0.8571	0.92105	0.96682	1
Eigen value	3.428433	0.255759	0.183105	0.132703

The result of PC_1 of this study was higher than that reported by Putra *et al* [16] that PC_1 explained 47.89% of total variance in body measurements of Pasundan cows, reported by Khargharia et al [17] that PC_1 of morphological traits was 40.37% of total variation in predicting BW of Assam Hill goat and that reported by Canaza-Cayo *et al* [14] that PC_1 of body measurements was 33.67% of total variation to predict BW of Corriedale ewes. De Campos *et al* [18] reported that PC_1 of morpho structural traits explained 57.70% of total variance in predicting BW of West African Dwarf sheep.

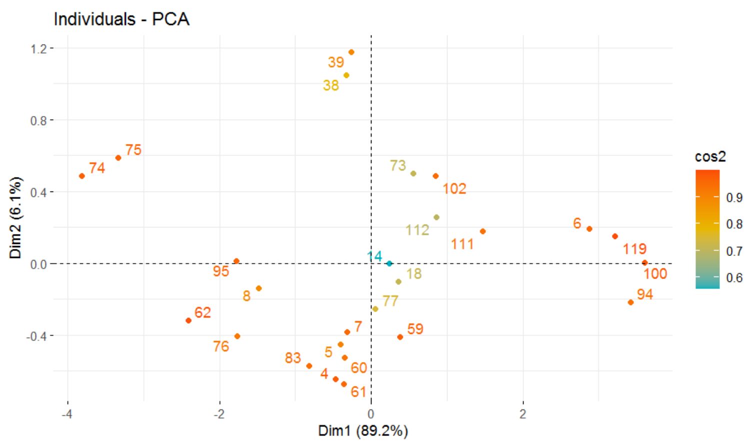


Figure 2. Individual PCA of body measurements in male OG cattle

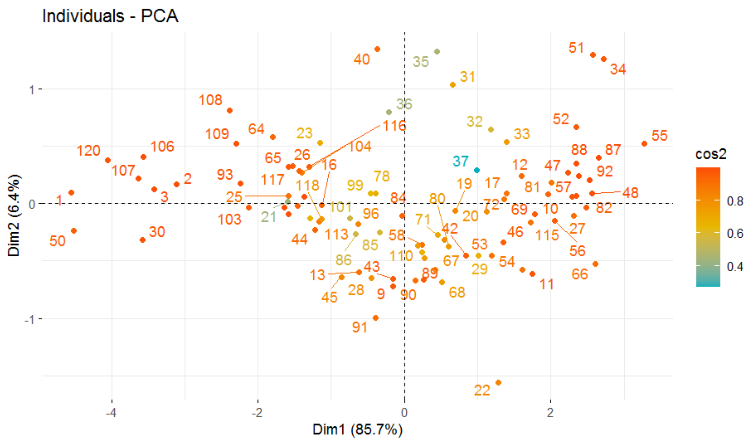


Figure 3. Individual PCA of body measurements in female OG cattle

3.3. Regression models for predicting body weight of Ongole Grade cattle

Regression models for predicting body weight of OG cattle using PC generated from body measurements and using original data of body measurements are presented in Table 3. The results of this study indicated that BW prediction using 4 PCs resulted in higher R^2, AIC, and BIC, but lower RSE and adjusted R^2 in male cattle compared to if using 3 PCs. However, regression coefficient for the 4th PC was not significant (P<0.05). This suggested that using 3 PCs is better because of higher adjusted R^2 and lower AIC and BIC. In addition, ANOVA test of the two models resulted in non-significant differences (P>0.05) meaning that using 3 PCs is more efficient and parsimonious than using 4 PCs. The same condition in female cattle, that using 2 PCs is better and more parsimonious for predicting BW of OG cattle. Based on Table 3 the regression model using 3 PCs generated from body measurements for predicting BW of male OG cattle is BW = 316.192 + 33.105*PC_1 + 13.381*PC_2 + 21.507*PC_3.

Body weight prediction using original body measurements of male OG cattle indicated that the full model (BW = -353.8314 + 2.7280*CG + 1.227*BL - 0.4281*SH + 3.44*CW) resulted in higher R^2, RSE, AIC, and BIC but lower adjusted R^2 compared to reduced model resulted from step-wise regression method (BW = -328.6121 + 3.1438*CG + 3.6361*CW). This result suggested that the later regression model is better and more parsimonious. This happened might be because of non-significant regression coefficient of BL and SH (Table 3). On the other hand, the full regression model and regression model resulted from step-wise regression method in female OG cattle was the same. This suggested that the use of full regression model is better.

Table 3. Regression model using PCA generated from body measurements and using original data of body measurements for predicting body weight of male and female OG cattle along with their selection criterion

Using PCA									
Intercept	Regression coefficient				R^2	Adj.R^2	RSE	AIC	BIC
	PC_1	PC_2	PC_3	PC_4					
Male									
316.192	33.105**	13.381**	21.51**	-14.86^{ns}	0.9655	0.959	13.09	213.949	221.498
316.192	33.105**	13.381**	21.51**		0.9629	0.9579	13.26	213.846	220.137
316.192	33.105**	13.381**			0.9476	0.9430	15.42	220.853	225.885
316.192	33.105**				0.9371	0.9345	16.53	223.587	227.361

Female									
314.766	24.822**	10.134**	5.639^{ns}	4.9788^{ns}	0.9164	0.9127	14.3	773.815	789.075
314.766	24.822**	10.134**	5.639^{ns}		0.9150	0.9122	14.34	773.381	786.098
314.766	24.822**	10.134**			0.9125	0.9106	14.47	774.090	784.264
314.766	24.822**				0.9013	0.9003	15.29	783.426	791.056

Using original data

Intercept	Regression coefficient				R^2	Adj.R^2	RSE	AIC	BIC
	CG	BL	SH	CW					
Male with full model (first raw) and using stepwise regression method (second raw)									
-353.831	2.7280**	1.2270^{ns}	-0.4281^{ns}	3.440**	0.966	0.959	13.09	213.95	221.498
-328.612	3.1438**			3.6361**	0.964	0.961	12.78	211.10	216.135
Female with full model (first raw) and using stepwise regression method (second raw)									
-442.619	1.7369**	1.5433**	1.1296^{ns}	3.966**	0.9164	0.913	14.3	773.82	789.075
-442.619	1.7369**	1.5433**	1.1296^{ns}	3.966**	0.9164	0.913	14.3	773.82	789.075

Transformation from PCA to original data

Intercept	Regression coefficient				R^2	Adj.R^2	RSE	AIC	BIC
	CG	BL	SH	CW					
Male									
0.00000	0.5094350	0.5163431	0.4941875	0.4792124	0.947				
Female									
0.00000	0.5032753	0.5064179	0.4894263	0.5007162	0.905				

In comparison between the use of regression model using principal component generated from body measurement and using original data of body measurements in predicting BW of OG cattle, the result of this study indicated that both the full and reduced models resulted in the same selection criterion. However, Pearson's correlation among body measurements (raw data) were quite high (>0.70) (Table 1) indicating that there was multicollinearity among the variables. It was supported by the high variance inflation factor (VIF) which was more than 5 (Table 5). While Pearson's correlation among PCs were almost zero (~0) (Table 4) indicating that there was no multicollinearity among the variables. It was supported by the low variance inflation factor (VIF) which was lower than 5 (Table 5). The result of this study suggested that the use of principal component generated from body measurements is better than the use of original body measurements in predicting BW of OG cattle.

Table 4. Pearson's correlations among PCs of body measurements of male OG cattle (above diagonal) and female OG cattle (below diagonal)

	PC_1	PC_2	PC_3	PC_4
PC_1	1	3.06E-16	1.25E-15	-5.36E-16
PC_2	3.22E-16	1	-6.40E-16	1.08E-15
PC_3	8.51E-16	-7.94E-16	1	-2.63E-16
PC_4	-1.10E-15	4.45E-16	1.13E-16	1

Notes: PC1-4 = principal component 1-4

Table 5. Variance inflation factor (VIF) among body measurements and among PCs

Sex	Using body measurements			
	CG	BL	SH	CW
Male	9.832032	13.4416	5.143783	3.463664
Female	4.3649	4.846024	3.476515	4.511088
	Using PCA			
	PCA_1	PCA_2	PCA_3	PCA_4
Male	1	1	1	1
Female	1	1	1	1

Notes: CG = chest girth, BL = body length, SH = shoulder height, CW = chest girth, PC1-4 = principal component 1-4.

When conducting a regression analysis with PCA scores as predictor variables, it can prove to be a complex task to make sense of the outcomes in relation to the original variables' meanings. This challenge arises because PCA scores are essentially linear combinations of the original variables and lack a straightforward, direct association with the actual measurements.

To tackle this dilemma, it frequently becomes necessary to revert the PCA scores to their original data context (Table 3). This reversion process involves using the loadings of the principal components, which capture the connections between the original variables and the PCA scores. Through this reverse transformation, we can reframe the results of the regression analysis within the context of the original variables. This, in turn, enables a more meaningful interpretation of the relationship between the predictors and the target variable.

Using PCA scores as predictors in a regression model can offer benefits such as dimensionality reduction and multicollinearity reduction, but the interpretability of the results may be challenging. To gain a better understanding of the relationships between the variables, it is essential to transform the PCA scores back to the original data space, allowing for a more insightful discussion and interpretation of the regression results. So the final regression model utilizing PC_1 transformation for predicting male and female OG cattle weight would be BW = 0.000 + 0.509435*CG + 0.5163431*BL + 0.4941875*SH + 0.4792124*CW and BW = 0.000 + 0.5032753*CG + 0.5064179*BL + 0.4894263*SH + 0.5007162*CW with R^2 of 0.947 and 0.905, respectively.

4. Conclusion

In conclusion, variables for body measurements were represented mostly by PC_1 for both in male (89.21%) and female (85.71%) OG cattle. Regression models developed using scores derived from one PC was the most efficient and parsimonious although using three PCs which explained 98.77% of total variation in BW of male OG cattle and two PCs which explained 92.11% of total variation in BW of female OG cattle are better in terms of higher R^2 for predicting BW with no multicollinearity problem. While both conventional linear regression and PCA-derived regression equations produce similar selection criteria, the superiority of PCA-derived regression equations lies in their lack of multicollinearity.

Acknowledgments

The authors express their deep appreciation to the Institute for Research and Community Service (LPPM), University of Lampung, for their generous funding of this research under the Professorship Research Scheme for the year 2022, with contract number: 473/UN26.21/PN/2022.

Animal Ethics and Welfare

All animal procedures related to data collections of Ongole Grade cattle were conducted in accordance with the Institute for Research and Community Service (LPPM), University of Lampung No. 473/UN26.21/PN/2022.

References

[1] Dakhlan A, Saputra A, Hamdani MDI and Sulastri 2020 Regression Models and Correlation Analysis for Predicting Body Weight of Female Ettawa Grade Goat using its Body Measurements *Adv. Anim. Vet. Sci.* **8(11)** 1142–1146 https://doi.org/10.17582/journal.aavs/2020/8.11.1142.1146
[2] Dakhlan A, Hamdani MDI, Putri DR, Sulastri and Qisthon A 2021 Short communication: Prediction of body weight based on body measurements in female saburai goat. *Biodiversitas.* **22(3)** 1391–1396 https://doi.org/10.13057/biodiv/d220341
[3] Dakhlan A, Qisthon A and Hamdani MDI 2021 Predicting Body Weight Based on Body Measurements at Different Ages and Sex in Saburai Goat *Adv. Anim. Vet. Sci.* **9(11)** 1791–1799 https://doi.org/10.17582/journal.aavs/2021/9.11.1791.1799
[4] Afolayan RA, Adeyinka IA and Lakpini CAM 2006. The estimation of live weight from body measurements in Yankasa sheep *Czech J. Anim. Sci.* **51(8)** 343–348
[5] Ibrahim A, Artama WT, Budisatria IGS, Yuniawan R, Atmoko BA and Widayanti R 2021 Regression model analysis for prediction of body weight from body measurements in female batur sheep of Banjarnegara district, Indonesia *Biodiversitas.* **22(7)** 2723–2730 https://doi.org/10.13057/biodiv/d220721
[6] Sabbioni A, Beretti V, Superchi P and Ablondi M 2020 Body weight estimation from body measures in Cornigliese sheep breed *Ital. J. Anim. Sci.* **19(1)** 25–30 https://doi.org/10.1080/1828051X.2019.1689189
[7] Budianto D, Widi TSM, Panjono P, Budisatria IGS and Hartatik T 2022. Estimation of Body Weight Using Linear Body Measurements in Two Crossbred Beef Cattle *Proc. 9th Int. Seminar on Tropical Anim. Prod. (ISTAP 2021)* Vol 18 (Yogyakarta, Indonesia) p 332–337 https://doi.org/10.2991/absr.k.220207.070
[8] Agung PP, Putra WBP, Anwar S and Wulandari AS 2018 Body Weight Estimation of Bali Cattle in Banyumulek Techno Park, West Nusa Tenggara using Several Morphometric Parameters *Bul. Peternak.* **42(1)** 20–25 https://doi.org/10.21059/buletinpeternak.v42i1.29840
[9] Paputungan U, Hakim L, Ciptadi G and Lapian HFN 2013 The estimation accuracy of live weight from metric body measurements in Ongole grade cows *J. Indones. Trop. Anim. Agric.* **38(3)** 149–155 https://doi.org/10.14710/jitaa.38.3.149-155
[10] Paputungan U, Hakim L, Ciptadi G and Lapian HFN 2015 Application of body volume formula for predicting live weight in Ongole crossbred cows *Int. J. Livest. Prod.* **6(3)** 35–40 https://doi.org/10.5897/ijlp2014.0243
[11] Putra WPB 2020 The Assesment of Body Weight of Sumba Ongole Cattle (Bos indicus) by Body Measurements. Manas *J. Agric. Vet. Life Sci.* **10(1)** 52–57
[12] Shoimah US, Dakhlan A, Sulastri and Hamdani MDI 2021 Use of body measurements to predict live body weight of Simmental bull in Lembang Artificial Insemination Center, West Java *IOP Conference Series: Earth and Environ. Sci.* **888** (Padang, Indonesia) p 012030. https://doi.org/10.1088/1755-1315/888/1/012030
[13] Negash, F 2021 Application of principal component analysis for predicting body weight of Ethiopian indigenous chicken populations *Trop. Anim. Health Prod.* **53** 104 https://doi.org/10.1007/s11250-020-02526-w
[14] Canaza-Cayo AW, Mota RR, Amarilho-Silveira F, Duarte AS and Cobuci JA 2021. Principal Component Analysis for Body Weight Prediction of Corriedale Ewes from Southern Peru *J. Anim. Heal. Prod.* **9(4)** 417–424 https://doi.org/10.17582/journal.jahp/2021/9.4.417.424

[15] R Core Team 2020 R: A Language and Environment for Statistical Computing, Vienna, Austria. https://www.r-project.org/
[16] Putra WPB, Said S and Arifin J 2020 Principal component analysis (PCA) of body measurements and body indices in the Pasundan cows *Black Sea J. Agric.* **3(1)** 49–55
[17] Khargharia G, Kadirvel G, Kumar S, Doley S, Bharti PK and Das M 2015 Principal component analysis of morphological traits of Assam hill goat in eastern Himalayan India *J. Anim. Plant Sci.* **25(5)** 1251–1258
[18] De Campos JS, Ikeobi CON, Olowofeso O and Smith OF 2014 Multivariate principal components analysis of the morphostructural traits of West African Dwarf sheep *Niger J. Anim. Prod.* **41(2)**:34–43 https://doi.org/10.51791/njap.v41i2.771