Journal of Agriculture and Food Environment 3 (2): 92-101; 2016

Application of Principal Component Analysis to Morphological Traits of Grass cutters (*Thryonomys swinderianus)*

I. Udeh

Department of Animal Science, Delta State University, Asaba Campus,
Asaba, Nigeria.

drudeh2005@yahoo.com

Date submitted 04/05/16

ABSTRACT

This study was designed to provide an objective description of the relationship between bodyweight and body measurements in grass cutters and to predict bodyweight from their orthogonal traits using principal component analysis (PCA). Bodyweight and seven body measurements namely: body length (BL), heart girth (HG), height at withers (HW), head length (HD), forelimb length (FL), hind limb length (HL) and tail length (TL) were measured in 104 female grass cutters at 4 – 6 months and 6 – 8 months of age, respectively. Phenotypic correlations among bodyweight and body measurements ranged from 0.01 – 0.79 at 4 – 6 months and from 0.09 – 0.84 at 6 – 8 months of age, the coefficients were mostly significant (p<0.05, 0.01) in both age periods. Two principal components were extracted after varimax rotation in each age period which explained 68.38% (4 – 6 months) and 81.02% (6 – 8 months) of the total variance. The first principal component (PC1) defined the body dimension of grass cutters at 4- 6 months of age while the second principal component (PC2) the extremity (head length and tail length). Similarly, in 6 – 8 months of age, PC1 defined the extremity while PC2 was mostly associated with body length, heart girth and tail length. The principal component based regression models predicted bodyweight of grass cutters with higher accuracy (R^2 = 62.9 – 83.5%) compared with interdependent based regression model (R^2 = 59.50 – 80.10%). The information obtained in this study could be applied for characterization and genetic improvement of grass cutters through selection.

KEYWORDS: Body measurements, bodyweight, grass cutters, multi-collinearity, stepwise multiple regression

INTRODUCTION

Grass cutters play important role in West Africa as source of meat and income. Its meat is a local delicacy in West Africa. Information on the relationships between bodyweight and body measurements in grass cutters is scant. The linear body measurements of meat animals such as the grass cutters give an indication of the bodyweight gain, skeletal size and carcass characteristics (Ebegbulem, 2012). Linear body measurements differ according to the factors such as breed, gender, yield and age (Pesmen and Yardimci, 2008). Estimation of bodyweights using body measurements is done using different statistical methods (Gurcan, 2000). This is very useful in areas where measuring scales are lacking and objective assessment of the animal is needed. Correlation and multiple regression analysis are used to study the complex relationships among biometrical traits. The correlations among body measurements may be different if the body measurements are treated as bivariates rather than multivariate during analysis (Shahin and Hassan, 2000). This is because of the interrelatedness or multicollinearity of the explanatory variables (Eyduran *et al*, 2010; Ogah, 2011). Principal component analysis and ridge regression are the two statistical procedures used to reduce or eliminate the problem of multicollinearity among predictor variables (Malau-Aduli *et al*, 2004). The central idea of principal component analysis is to reduce the dimensionality of a data set consisting of a large number of interrelated variables, while retaining as much as possible of the variation present in the data set. This is achieved by transforming to a new set of variables, the principal components which are ordered so that the first few retain most of the variation present in all of the original variables (Jolliffe, 1986). PCA is a powerful tool to analyze data and once patterns in the data are found, it can be compressed by reducing the number of dimensions without much loss of information (Sayad, 2010). In recent times, there are lots of researches on the domestic grass cutters. This is because of the prospect attached to grass cutter farming as a means of bridging the gap of animal protein shortages in West Africa. Most of the researches on grass cutters focused on the husbandry, biology and genetics of the animal. Information on the complex relationship among morphological traits of the animal is very scanty. The present study was aimed at understanding the relationship between bodyweight and linear body measurements in grass cutters using multivariate principal component analysis.

MATERIALS AND METHODS

The experiment was conducted at the grass cutter research unit of the Department of Forestry and Wildlife, Delta State University, Asaba Campus, Asaba, Delta State, Nigeria. Asaba is located on latitude 60^0 12' North and longitude 60^0 45' East. Annual rainfall in Asaba ranges from 1800 to 3000mm. Maximum temperature in the area is from 27.5^0C to 39.9^0C (Federal Ministry of Aviation, Meteorological Services, Asaba).

A total of 104 female grass cutters from 26 families were used, each family being made up of 4 females and 1 male. The grass cutters were procured at 4 months of age from a reputable farm at Amukpe, Delta State. The grass cutters were housed in concrete cages with each family occupying a cage. Each cage was partitioned into two compartments (a room and parlour). The dimension of the room was 1.00m (length), 1.00m (width) and 1.00m (height). The dimension of the parlour was the same as the room except that the width was 0.9m. The entrance of the cage was on top and made of wooden door for room and wire mesh for parlour. The concrete cages were constructed in a solid building with dwarf walls. The roof of the building was constructed with wood and asbestos. The animals were fed with grasses such as guinea grass (*Panicum maximum*) and elephant grass (*Pennisetum purpureum*). The grasses were supplemented with maize and soya bean concentrates. Clean drinking water was also provided. Both feed and water were made available all the time.

The bodyweights of the grass cutters were taken on weekly basis from 4 to 8 months of age using a weighing balance and recorded in kilograms. The linear body measurements (cm) were obtained using a tape. All the measurements were taken in the morning from 7 am to 10 am before feeding. The following linear body measurements as described by Annor *et al* (2011) and Ogbuewu *et al* (2010) were taken:

Body length: The distance from the tip of the nose to the tip of the tail.

Head length: Measured as the distance from the tip of the nose to the junction of the head and the neck.

Heart girth: This was determined by measuring the circumference of the chest directly below the arms.

Height at withers: The distance from the surface of a platform to the withers.

Tail length: Distance from base of tail to the tip.

Forelimb length and hind limb length: Measured from the shoulder and pelvic joints to the tip of the paws respectively.

Descriptive statistics such as mean, standard deviation and coefficient of variation were calculated for bodyweight and each linear body measurement. Pearson's coefficients of correlation between bodyweight and linear body measurements were calculated. The suitabilityor appropriateness of the data to principal component analysis was tested using Kaiser Meyer Olkin (KMO) measure of sample adequacy and Bartlett's test of sphericity. Anti-image correlations from the correlation matrix were examined to further confirm the appropriateness of the data to principal component analysis to which the data were subjected. Bodyweight of grass cutters was predicted from the linear body measurements and the factor scores using stepwise multiple regression analysis. The following stepwise regression models described by Ogah (2011) were used:

$BW = a + B_iX_i + ... + B_KX_K$.................... (a)

$BW = a + B_iPC_i + ... + B_KPC_K$.................. (b)

Where,

BW is the bodyweight,

a is intercept,

B_i is the i^{th} partial regression coefficient of the i^{th} linear body measurement;

X_i or the i^{th} principal component (PC).

The principal component analysis was performed using the factor programme of SPSS 16 (2007).

RESULTS AND DISCUSSION

Means, standard deviations and coefficient of variations of bodyweight and linear body measurements of grass cutters at 4 – 6 months and at 6 – 8 months of age are presented in Table 1. The mean bodyweight of grass cutters at 4 – 6 months and at 6 – 8 months were 0.75 kg and 1.43 kg respectively. Ikpeze and Ebenebe (2004) reported bodyweight range of 0.66 – 2.51 kg for grass cutters 2 – 10 months of age. Annor *et al* (2011) reported range of 0.69 – 1.94 kg for male and female grass cutters aged 4 – 5 months.The discrepancy in the mean bodyweight of grass cutters reported in this study and the range reported by others may be attributed to environmental differences. The values of mean body measurements obtained in this study were similar to those reported by Udeh and Okonta (2013). The highest variability was shown by bodyweight in both age periods. The coefficient of variations of body measurements were low and ranged from 6.18 – 10.59% at 4 – 6 months and from 4.32 – 7.46% at 6 – 8 months of age.

Table 1: Descriptive statistics of bodyweight (kg) and linear body measurements (cm) of grass cutters.

Traits	4-6months			6-8months		
	Mean	**SD**	**CV**	**Mean**	**SD**	**CV**
Body Weight	0.75	0.15	20.00	1.43	0.33	23.08
Body Length	42.89	2.65	6.18	47.58	3.55	7.46
Heart Girth	22.59	1.46	6.46	25.02	1.08	4.32
Height at Withers	12.71	0.93	7.32	14.63	0.83	5.67
Head Length	8.57	0.54	6.30	9.73	0.57	5.86
Forelimb Length	7.84	0.83	10.59	9.39	0.52	5.54
Hind limb Length	13.01	1.01	7.76	14.65	0.77	5.26
Tail Length	13.57	0.94	6.93	15.26	1.13	7.40

SD – standard deviation, CV – coefficient of variation (%).

Table 2: Pearson correlation coefficients among body weight and linear body measurements of grass cutters at 4 – 6 months (lower matrix) and 6 – 8 months (upper matrix)

Traits	BW	BL	HG	HW	HD	FL	HL	TL
BW		0.45**	0.67**	0.77**	0.75**	0.68**	0.77**	0.62**
BL	0.74**		0.62**	0.51**	0.39**	0.09	0.56**	0.74**
HG	0.79**	0.50**		0.62**	0.61**	0.43**	0.82**	0.71**
HW	0.65**	0.58**	0.66**		0.72**	0.54**	0.66**	0.84**
HD	0.25*	0.21	0.27*	0.19		0.67**	0.67**	0.55**
FL	0.64**	0.47**	0.58**	0.58**	0.01		0.60**	0.29*
HL	0.73**	0.69**	0.63**	0.78**	0.17	0.56**		0.64**
TL	0.07	0.22	0.13	0.19	0.16	-0.25	0.10	

*BW – body length, BL – body length, HG – heart girth, HW – height at withers, HD – head length, FL – fore limb length, HL – hind limb length, TL – tail length, *correlation is significant at the 0.05 level (2-tailed), **correlation is significant at 0.01 level (2-tailed).*

It can be inferred from this study that the body measurements tend to be more uniform compared to bodyweights of grass cutters. Pearson's coefficients of correlation among the variables are shown in Table 2. The coefficient of correlations among the variables ranged from 0.01 – 0.79 at 4 – 6 months and 0.09 – 0.84 at 6 – 8 months of age and were mostly significant (P<0.05, 0.01). This indicates high predictability among the variables. The highest correlation coefficient was bodyweight and heart girth (r =0.79) at 4 – 6 months and height at withers and tail length (r = 0.84) at 6 – 8 months of age. Annor *et al* (2011) reported high phenotypic correlations (0.71 – 0.90) between bodyweight and morphological traits in grass cutters. Similar results were also reported in grass cutters (Ikpeze and Ebenebe 2004). The calculated value of KMO for grass cutters aged 4 – 6 months and at 6 – 8 months were 0.807 and 0.837 respectively, thus indicating that the sample size was adequate for PCA. According to Field (2000), a sample is adequate if the value of KMO is greater than 0.50.

Furthermore, all the elements in the diagonal of the anti-image correlation were greater than 0.50 at both age periods thereby confirming that the sample sizes were adequate. Bartlett's test of sphericity tests the null hypothesis that the original correlation matrix is an identity matrix which would indicate that the PCA or factor model is appropriate.

Table 3: Eigen values and share of total variance along with factor loadings after varimax rotation and communalities of the body measurements of grass cutters aged 4 - 6 months and 6 - 8 months

	4 – 6 months			6 – 8 months		
Traits	**PC1**	**PC2**	**Communality**	**PC1**	**PC2**	**Communality**
Body Weight	0.90	0.12	0.83	0.79	0.45	0.84
Body Length	0.76	0.26	0.65	0.04	0.92	0.84
Heart Girth	0.82	0.17	0.70	0.49	0.72	0.75
Height at Withers	0.83	0.17	0.72	0.62	0.62	0.77
Head Length	0.15	0.71	0.52	0.80	0.36	0.78
Forelimb Length	0.79	-0.22	0.68	0.93	-0.02	0.87
Hind limb Length	0.87	0.10	0.77	0.66	0.59	0.78
Tail Length	0.03	0.78	0.60	0.31	0.88	0.86
Eigen value	4.30	1.168		5.30	1.19	
% of total variance	53.78	14.59		66.20	14.82	

The estimates of Bartlett's test of sphericity obtained in this study at 4 – 6 months (chi-square =359.16, P<0.001) and 6 – 8 months (chi-square = 639.54, p<0.001) were significant, thus confirming that PCA was appropriate for the data. Table 3 is a summary of the principal component analysis showing the components extracted after varimax rotation (PC1 and PC2), eigen values, share of total variance and communalities.

Communalities represent the proportion of variance in the original variable shared in common with all other variables included in the analysis. The communalities ranged from 0.52 to 0.83 (4 – 6 months) and from 0.75 – 0.87 (6 – 8 months). This implies that the variance of each variable was well reflected through the extracted components and hence PCA adequate. Two principal components from each age period with eigen value greater than 1 were extracted after varimax rotation, accounting for 68.38% (4 – 6 months) and 81.02% (6 – 8 months) of the total variance present in the original variables. The principal component coefficients (factor loadings) show the relative contribution of each body measurements to a particular principal component (Mavule *et al.*, 2013). In 4 – 6 months, PC1 which accounted for 53.78% of the total variance loaded heavily on all the body measurements except HD and TL while PC2 with high positive loading for HD and TL explained 14.59% of the total variance. The implication is that PC1 defined the body dimension of grass cutters at 4 – 6 months of age while PC2 the extremity. In 6 – 8 months, PC1 explained 66.20% of the total variance and was characterized by high positive loadings for FL, HD, HL and HW. PC2 accounted for 14.82% of the total variance and was highly related to BL, TL, HG, HW and HL. This means that PC1 mostly defined the extremity of grass cutters at 6 – 8 months of age while PC2 defined BL, TL and HG. The component score coefficients matrix is presented in Table 4.

Table 4: Component score coefficient matrix

	4 – 6 months		**6 – 8 months**	
Traits	**PC1**	**PC2**	**PC1**	**PC2**
Body Weight	0.22	-0.02	0.25	-0.02
Body Length	0.16	0.12	-0.28	0.47
Heart Girth	0.19	0.04	0.01	0.22
Height at Withers	0.19	0.04	0.11	0.12
Head Length	-0.05	0.57	0.29	-0.07
Forelimb Length	0.23	-0.28	0.48	-0.32
Hind limb Length	0.21	-0.02	0.14	0.10
Tail Length	-0.10	0.64	-0.13	0.36

These components were used to obtain principal component factor scores for regression analysis. Table 5 presents the stepwise multiple regression of bodyweight on the original body measurements of grass cutters and on their principal component factor scores at 4 – 6 months and 6 – 8 months of age. HG alone captured 62.70% of the variability in the bodyweight of grass cutters at 4 – 6 months of age. When BL was added to the model, the accuracy of bodyweight prediction in grass cutters (4 – 6 months) increased to 78.60%. The combination of HG, BL and FL in the regression equation increased further the accuracy of bodyweight prediction in grass cutters (4 – 6 months) to 80.10%. Annor *et al* (2011) reported that HG was the best predictor of body weight of grass cutters, followed by BL.

Table 5: Stepwise multiple regression of bodyweight on original body measurements and their principal component factor scores

Variables (4 – 6 months)	Models	R^2	S.E
Original body measurements			
as predictors		0.629	0.094
1 HG	BW = -1.129 + 0.083 HG		
2 HG, BL	BW = -1.723 + 0.059 HG + 0.027 BL	0.786	0.072
3 HG, BL, FL	BW = -1.687 + 0.051 HG + 0.025BL + 0.029FL	0.801	0.070
Principal components as predictors.			
1 PC1	BW = 0.755 + 0.138 PC1	0.814	0.066
2 PC1, PC2	BW = 0.755 + 0.138PC1 + 0.018PC2	0.829	0.064
6 – 8 months			
Original body measurements			
as predictors		0.595	0.210
1 HL	BW = - 3.383 + 0.329HL		
2 HL, HW	BW = -4.141 + 0.201HL + 0.179HW	0.713	0.177
3 HL, HW, FL	BW = -4.563 + 0.155HL + 0.155HW + 0.154FL	0.749	0.167
Principal components as predictors			
1 PC1	BW = 1.431 + 0.260PC1	0.629	0.200
2 PC1, PC2	BW = 1.431 + 0.260PC1 + 0.148PC2	0.835	0.135

Similarly, Udeh and Okonta (2013) reported that TL (R^2=98%) and HG (R^2=97%) were better predictors of bodyweight of grass cutters compared to other body measurements. Generally, it has been established that HG is a very good predictor of body weight in domestic animals such as cattle (Abdelhadi and Babiker, 2009), goat (Yakubu *et al*, 2011; Chitra *et al*, 2012), sheep (Oke and Ogbonnaya, 2011; Mavule *et al*, 2013) and rabbits (Oke *et al*, 2011; Udeh, 2013). According to Okpeku *et al* (2011), the importance of HG in weight estimation could be as a result of the fat that muscle, some fat along with bone structure contributes to its formation. HL (R^2=59.50%) was the highest single predictor of bodyweight of grass cutters at 6 – 8 months of age. Inclusion of HW to the regression model increased the accuracy of bodyweight prediction in grass cutters (6 – 8 months) to 71.30%. The best prediction equation (R^2=74.90%) for the prediction of body weight of grass cutters (6 – 8 months) was given by the combination of HL, HW and FL. However, one has to be cautious when using interdependent variables such as heart girth and body length as predictors of bodyweight in animals because multicollinearity was shown to be associated with unstable estimates of regression coefficients which could lead to unreliable predictions (Malau-

Aduli *et al*, 2004). The use of orthogonal traits (principal components) derived from the original body measurements are more reliable in weight estimation. As shown in Table 5, PC1 predicted bodyweight of grass cutters at 4 – 6 months and 6 – 8 months of age with R^2 = 81.40% and 62.90% respectively. A combination of PC1 and PC2 in the regression model increased the accuracy of bodyweight prediction in grass cutters to 82.90% (4 – 6 months) and 83.90% (6 – 8 months) of age. In both age periods, the accuracy of bodyweight prediction increased when orthogonal traits were used compared to interdependent traits. Similar observations have been reported in sheep (Mavule *et al*, 2013) and in two species of fish (Yakubu *et al.*, 2012).

CONCLUSION

Principal component analysis was used to condense eight linear body measurements taken from 104 female grass cutters at 4 – 6 months and 6 – 8 months into two principal components each that describes their body conformation. The principal component factor scores predicted bodyweight of grass cutters with higher degree of accuracy compared to the original body measurements and at the same time eliminating multicollinearity which could have otherwise given unreliable prediction results and false inference. The results of this study suggest that PCA could be used in grass cutter breeding programmes to select animals based on group of variables rather than on isolated traits.

REFERENCES

Abdelhadi, O. M. A and Babiker, S. A. (2009).Prediction of zebu cattle live weight using live animal measurements. Livestock Research for RuralDevelopment. Vol. 21, article #133. http://www.Irrd21/8/abde21133.

Annor, S. Y., Ahunu, B. K., Aboagye, G. S., Boa-Amponsem, K., Djang-Fordjour, K. T and Cassady, J. P. (2011). The genetics of morphological traits in the grass cutters. Livestock Research for Rural Development vol. 23 article #167. http://www.Irrd.org/Irrd23/8/anno23167.htm

Chitra, R; Rajendran, S; Prasanna, D and Kirubakaran, A. (2012). Prediction of bodyweight using appropriate regression model in adult female Malabari goat. Veterinary world vol. 5 (7): 409-411.

Ebegbulem, V. N. (2012). Body conformation characteristics of domestic rabbits in humid tropical southern Nigeria. J of Social Science and Public Policy 4: 18-23.

Eyduran, E., Topal, M and Sonmez, A. Y. (2010). Use of factor scores in multiple regression analysis for estimation of body weight by several body measurements in brown trouts (*salmo trutta fario*). Int. J. Agric. Bio., 12: 611-615.

Field, A. (2000). Discovering statistics using SPSS for Windows. London – Thousand Oaks – New Delhi. Sage publications.

Gurcan, I. S. (2000). Estimation of live weight by statistical methods using body measurements in merino sheep. Ank. Univ. Health Sci. Inst. Ph.D thesis Ankara (in Turkish, English abstract).

Ikpeze, O. O and Ebenebe, C. I. (2004). Relationships between physical body traits of the grass cutter (Rodentia: Thryonomyidae) in Akpaka forest reserve, Onitsha. Anim. Res. Int. 1 (3): 160-163.

Jolliffe,I.T. (1986). Principal component analysis. Second ed. Springer. Page 32.

Malau-Aduli, A. E. O., Aziz, M. A., Kojima, T., Niibayashi, T., Oshima, K and Komatsu, M. (2004). Fixing collinearity instability using principal component and ridge regression analysis in the relationship between body measurements and body weight in Japanese black cattle. J. Anim. and Vet. Advances 3: 856-863.

Mavule, B. S; Muchenje, V; Bezuidenhout, C. C and Kunene, N. W. (2013). Morphological structure of zulu sheep based on principle component analysis of body measurements. Small Ruminant research 111: 23 – 30.

Ogah, D. M. (2011). Assessing size and conformation of the body of Nigerian indigenous turkey. Slovak J. Anim. Sci. 44 (1): 21-27.

Ogbuewu, I. P., Uchegbu, M. C., Okoli, I. C and Iloeje, M. U. (2010). Assessment of blood chemistry, weight gain and linear body measurements of pre-pubertal buck rabbits fed different levels of neem meals. Chilean J. Agric. Res. 70 (3): 515-520.

Oke, E. K and Ogbonnaya, E. O. (2011). Application of physical body traits in the assessment of breed and performance of WAD sheep in a humid tropical environment. Livestock Research for Rural Development. Vol. 23 article #024. http://www.Irrd.org/Irrd23/2/oke23024htm.

Oke, E. K; Herbet, U; Obike, O. M and Ogbonnaya, E. O. (2011). Effect of weaner body weight on growth traits of rabbits. Online Journal of Animal and Feed Research 1(1): 22-27.

Okpeku, M; Yakubu, A; Peters, S. O; Ozoje, M. O; Ikeobi, C. O. N; Adebambo, O. A and Imumorin, I. G. (2011). Application of multivariate principal component analysis to morphological characterization of indigenous goats in southern Nigeria. Acta Agriculture Slovenica, 98 (2): 101 – 109.

Pesman, G and Yardimci, Y. (2008). Estimating the live weight using some body measurements in Saanen goats. Arch. Zootechnica 11: 30-40

Sayad, S. (2010). Principal component analysis. http://chem-eng.utoronto.ca/~datamining/1:26.

Shahin, K. A and Hassan, N. S. (2000). Sources of shared variability among body shape characters at marketing age in New Zealand White and Egyptian rabbit breeds. Ann. Zootec. vol. 49(5): 435-445.

SPSS (2007). Statistical package for the social sciences. SPSS Inc. 444 Michigan avenue Chicago, IL 60611, USA.

Udeh, I. (2013). Prediction of body weight in rabbits using principal component factor scores in multiple linear regression model. Rabbit Genetics 3(1): 1-6

Udeh, I and Okonta, B. O. (2013). Prediction of bodyweight from body measurements in grass cutters. Trop. Anim. Health and Prod. Vol 45 (6): 1331-1334.

Yakubu, A., Okunsebor, S. A., Kigbu, A. A., Sotolu, A. O and Imgbian, T. D. (2012). Use of factor scores for predicting body weight from some morphometric measurements of two fish species in Nigeria. J. Agric. Sci. 4(8): 60-64.