RESEARCH ARTICLE

Principal Component Analysis of Biometric Traits Explain the Body Weight of HF Crossbred Cattle

Ashwini J. Patel*, Dharamsi N. Rank

Abstract

The present research was carried out to define the biometric traits of HF crossbred cattle from three farms around Anand (India). The biometric traits recorded on 506 cattle were body length (BL), height at wither (HW), height at hip (HH), heart girth (HG), chest depth (CD) and width of hip (WH), live body weight (BW) and age. All the data were grouped sex-wise. Principal component analyses of biometric traits revealed that out of six components, two principal components each were extracted in female and male group. The identified two components in female group could explain 95.88% of cumulative variance. First component accounted for 69.20 % of the variation. It was represented by significant positive high loading of BL (0.893), HW (0.911), HH (0.908), HG (0.879), CD (0.867) and WH (0.842). The second component explained 26.68 % of total variance with high loading of age (0.915). Out of two principal components, one (PC1) provided a means of reduction in the number of biometric traits to be recorded in HF crossbred female. In HF crossbred male group, the identified two components could explain 97.57% of cumulative variance. First component accounted for 60.58 % of the variation. It was represented by significant positive high loading of BL (0.836), HW (0.865), HH (0.901), HG (0.768), CD (0.804) and WH (0.737). The second component explained 36.99 % of total variance with high loading of age (0.888). First component seemed to be explaining the maximum of general body conformation in HF crossbred male. The result suggests that principal component analysis (PCA) could be used in breeding programs with a drastic reduction in the number of biometric traits to be recorded to explain body conformation and in selection of elite animals.

Key words: Biometric traits, Body weight, Factor analysis, HF crossbred cattle, Principal component analysis.

Ind J Vet Sci and Biotech (2023): 10.48165/ijvsbt.19.2.07

Introduction

Biometric traits are used to characterize the different breeds of livestock as they give an idea of body conformation. Biometric traits are also used for comparison of growth in different individuals. Body dimensions have been used to indicate breed, origin and relationship or shape and size of an individual. EAAP and FAO have used height at withers as a prime indicator for their type (Simon and Buchenauer, 1993). Recently, alternative body measurements and indices estimated from different combinations of different body traits produced a superior guide to weight and were also used as an indicator of type and function in domestic animals (Schwabe and Hall, 1989; Salako, 2006). Principal components analysis (PCA) technique by Hotelling (1933), is a multivariate ordination practice used to demonstrate arrangements in multivariate data. Linear combination with maximum variance is the first principal component (Johnson and Wichern, 2007). Morphometric variables are combined by this analytical tool to produce components or catalogue that are uncorrelated and data can be viewed from different dimensions (Manly, 1994). Analysis of variance and correlations are used to obtain relationships among different body measurements. The factor and PCA can explain relationships in a better way when the recorded traits are correlated. The purpose of PCA is to reduce a set of data that may describe and be used easily.

Department of Animal Genetics and Breeding, College of Veterinary Science and Animal Husbandry, Anand-388001, Kamdhenu University, Gujarat, India

Corresponding Author: Ashwini J. Patel, Department of Animal Genetics and Breeding, College of Veterinary Science and Animal Husbandry, Anand-388001, Kamdhenu University, Gujarat, India, e-mail: ashwini.patel108@gmail.com

How to cite this article: Patel, A.J. & Rank, D. N. (2023). Principal Component Analysis of Biometric Traits Explain the Body Weight of HF Crossbred Cattle. Ind J Vet Sci and Biotech. 19(2), 34-39.

Source of support: Nil

Conflict of interest: None

Submitted: 18/11/2022 **Accepted:** 12/12/2022 **Published:** 10/03/2023

For genetic improvement, principal components simultaneously consider a group of attributes which may be used for selection purpose. Fumio *et al.* (1982),Hammock and Shrode (1986), Karacaroen and Kadarmindeen (2008) and Yakuba *et al.* (2009) used factor analysis to study the different biometric traits in Japanese Black cattle, beef cattle, Swiss dairy cattle and White Flauni cattle, respectively. Salako (2006) and Sadek *et al.* (2006) used factor analysis to study the principal component factor analysis of the morpho-structural traits in Uda sheep and factor analysis of body measurements in Arabian horses, respectively. Presently, the size of the cow, represented by different body measurements, is one of the

important criteria in selection of elite animals. There is an urgent need to describe the body conformation by recording a minimum number of biometric traits which reduce the cost, labor and time. The present research was carried out to study the different body measurements, relationships among them, and to develop unobservable factors (latent) to define which of these measures best represent body conformation in HF crossbred cattle.

Materials and Methods

Data consisted of 6 different body measurements on HF crossbred cattle (both male and female) from Livestock Research Station, AAU, Anand; Sarsa Heifer Project - Amul Dairy, Anand; and Ode Semen Station – Amul Dairy, Anand, Gujarat, India. All measurements were recorded by the same recorder to avoid between-recorder effects. For body measurements simple tailor tap and sticks/scale, while for actual body weight platform weighing balance was used. During collection of body measurements, animals were made to stand on four legs squarely, relaxed/without stress. All measurements were in cm and body weight was measured in kg. The recorded body measurements were body length (BL), height at wither (HW), height at hip (HH), heart girth (HG), chest depth (CD) and width of hip (WH).

Actual body weight with exact age and all the above measurements were collected from three different farms. All the data were assorted sex-wise in male and female groups. Actual body weight of an animal was considered as dependent variable and all body measurements were considered as independent variables. Direct and indirect causal effects of all linear body measurements on body weight were established based on path analysis by SPAR1 software.

To estimate the body weight of the cattle, regression model used was as per equation (1).

$$Y= a+b_1X_1+b_2X_2+b_3X_3+b_4X_4+b_5X_5+b_6X_6 + E \ldots\ldots\ldots\ldots\ldots\ldots\ldots (1)$$

The model consists of one dependent variable; Y = body weight, and five independent variables; X_1 =body length, X_2 =height at withers, X_3 = height at hip, X_4 =heart girth, X_5= chest depth and X_6= width of hip where, a is intercept, b is regression coefficient and E is error.

PCA was carried out to find out which factor has highest effect on body weight by SPSS software. The PCA was applied to six body conformation traits and Kaisee-Meyor-Olkin (KMO) measure of sampling adequacy (MSA) was obtained. The estimate of sampling adequacy KMO revealed the proportion of the variance in different biometric traits caused by the underlying components (Kaiser, 1958). The overall significance of the correlation matrix was tested with Bertlett's test of sphericity for the biometric traits. The MSA below 0.5 was not accepted, as KMO-MSA greater than 0.5 is must for satisfactory factor analysis to proceed.

To decide how many components should be extracted, cumulative variance of components (up to 95%) was used. Scree plot was also used to decide the actual number of the components to be included for analysis, components having Eigenvalues up to the point (the value after which adding new factor doesn't make significant difference in variation here its 6.30) "bent of elbow" was considered. The widely used and accepted method of rotation (varimax) was applied for rotation of principal components through the transformation of the components to approximate a simple structure. Factor analysis assumes that a variable's variance can be decomposed into two parts. The first part is called common variance (communality factor) that is shared by other variables included in the model. The estimate of communality for each variable measures the proportion of variance of that variable explained by all the other components jointly. The second part is called specific variance (unique factor) as it is specific to a particular variable and includes the error variance. Factor analysis deals only with the common variance of the observed variables. It assumes that the unique variance represents a small and significant portion of the total variance.

Results and Discussion

The PCA was applied to six body measurements in HF crossbred cattle in both groups. Group wise scree plots depicted the various components which gave idea to decide the actual number of the components to be included for analysis, components having Eigen values up to the point "bent of elbow" in scree plot and components which explained variance minimum 4.00 % and cumulative variance around 94-96 % were extracted in groups of male and female. Varimax rotation is key feature to choose components in principal factor which are strongly correlated with certain variables and uncorrelated with other variables. Varimax rotation mainly affects lower variance components. So, lower variance components having values less than 0.70 were dropped out and were used to make principal component in groups.

For HF crossbred female where age is one of the factor, KMO measure of sampling adequacy (MSA) was obtained as 0.937. The overall significance of the correlation matrix tested with Bertlett's test of sphericity for the biometric traits (chi-square = 5901.552, $P<0.01$) was significant, it means correlation matrix is not an identity matrix and provided enough support for the validity of the factor analysis of data.

Out of six components, two principal components were extracted (Table 1). The identified two components could explain 95.88% of cumulative variance. First component accounted for 69.20 % of the variation. It was represented by significant positive high loading of BL (0.893), HW (0.911), HH (0.908), HG (0.879), CD (0.867) and WH (0.842). First component seemed to be explaining the maximum of general body conformation in HF crossbred female. The second component

explained 26.68 % of total variance with high loading of age (0.915). The communality ranged from 0.942 (WH) to 0.996 (age) and unique factors ranged from 0.004 - 0.058 for all six different biometric traits (Table 2). Based on Eigen value higher than one, only one component could be extracted based on Scree plot (Fig. 1) and it explained around 90.00% of total variance and no varimax rotation could be applied, this suggested that the use of principal component one (PC1) provided a means of reduction in the number of biometric traits to be recorded in HF crossbred female.

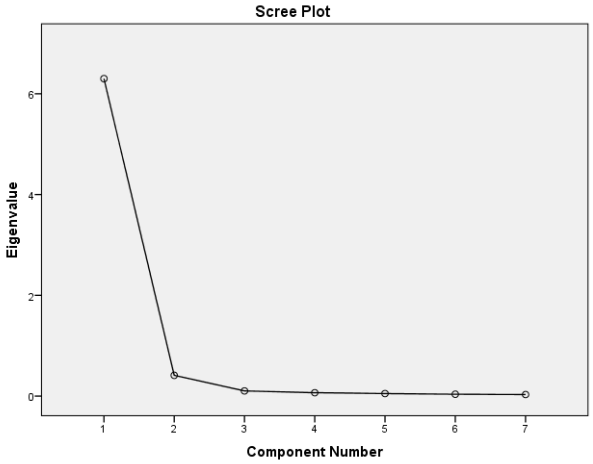

Fig. 1: Scree plot showing component number with Eigenvalue of HF crossbred female (including age factor)

For HF crossbred female (excluding age factor), KMO measure of sampling adequacy (MSA) was obtained as 0.925. The overall significance of the correlation matrix tested with Bertlett's test of sphericity for the biometric traits (chi-square = 5515.716, P<0.01) was significant, it means correlation matrix is not an identity matrix and provided enough support for the validity of the factor analysis of data.

Out of six components, two principal components extracted which could explain 96.89 % of cumulative variance (Table 3). First component accounted for 49.35% of the variation. It was represented by significant positive high loading of BL (0.713), HG (0.745), CD (0.741) and WH (0.836). The second component explained 47.54 % of total variance with high loading of BL (0.809) and HH (0.809). The communality ranged from 0.950 (HG) to 0.981 (CD) and unique factors ranged from 0.019 to 0.050 for all six different biometric traits (Table 4). Based on Eigenvalue higher than one, only one component could be extracted based on scree plot (Fig. 2) and it explained around 94.99 % of total variance and no varimax rotation could be applied.

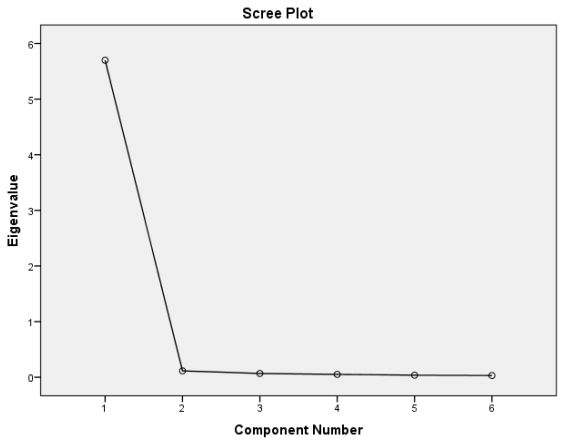

Fig. 2: Scree plot showing component numbers with Eigenvalue of HF crossbred female (excluding age factor)

Table 1: Total variance explained by different components in HF crossbred female (including age factor)

Component	Eigenvalue			Extraction Sums of Squared Loadings			Rotation Sums of Squared Loadings		
	Total	% Variance	Cumulative %	Total	% Variance	Cumulative %	Total	% Variance	Cumulative %
1	6.30	90.00	90.00	6.30	90.00	90.00	4.84	69.20	69.20
2	0.41	5.88	95.88	0.41	5.88	95.88	1.87	26.68	95.88
3	0.10	1.47	97.35	0.10	1.47	97.35	0.09	2.1	97.98
4	0.07	0.96	98.32	0.07	0.96	98.32	0.06	1.02	99
5	0.05	0.72	99.04	0.05	0.72	99.04	0.03	0.07	99.07
6	0.04	0.53	99.56	0.04	0.53	99.56	0.01	0.03	99.11

Table 2: Varimax rotated component matrix (a) of different factors for biometric traits with its communalities and unique factor in HF crossbred female (Component "a" means BW)

Traits	Component "a"		Communalities	Unique Factor
	1	2		
Age	0.399	0.915	0.996	0.004
BL	0.893	0.396	0.954	0.046
HW	0.911	0.354	0.955	0.045
HH	0.908	0.360	0.955	0.045
HG	0.879	0.438	0.964	0.036
CD	0.867	0.440	0.945	0.055
WH	0.842	0.483	0.942	0.058

Table 3: Total variance explained by different components in HF crossbred female (excluding age factor)

Component	Eigenvalue			Extraction Sums of Squared Loadings			Rotation Sums of Squared Loadings		
	Total	% Variance	Cumulative %	Total	% Variance	Cumulative %	Total	% Variance	Cumulative %
1	5.70	94.99	94.99	5.70	94.99	94.99	2.96	49.36	49.36
2	0.11	1.91	96.90	0.11	1.91	96.90	2.85	47.54	96.90
3	0.07	1.13	98.03	0.07	1.13	98.03	0.08	1.15	98.05
4	0.05	0.85	98.88	0.05	0.85	98.88	0.04	0.8	98.85
5	0.04	0.61	99.49	0.04	0.61	99.49	0.03	0.5	99.35
6	0.03	0.51	100.00	0.03	0.51	100.00	0.02	0.5	99.85

Table 4: Varimax rotated component matrix (a) of different factors for biometric traits with its communalities and unique factor in HF crossbred female (excluding age factor)

Traits	Component "a"		Communalities	Unique Factor
	1	2		
BL	0.713	0.668	0.955	0.045
HW	0.570	0.809	0.980	0.020
HH	0.571	0.809	0.980	0.020
HG	0.745	0.643	0.968	0.032
CD	0.741	0.633	0.950	0.050
WH	0.836	0.532	0.981	0.019

For HF crossbred male (including age factor), KMO measure of sampling adequacy was obtained as 0.913. The overall significance of the correlation matrix tested with Bertlett's test of sphericity for the biometric traits (chi-square = 1185.729, P<0.01) was significant means correlation matrix is not an identity matrix.

Out of six components in male, two principal components were extracted which could explain 97.57% of cumulative variance. First component accounted for 60.58 % of the variation (Table 5). It was represented by significant positive high loading of BL (0.836), HW (0.865), HH (0.901), HG (0.768), CD (0.804) and WH (0.737). First component seemed to be explaining the maximum of general body conformation in HF crossbred male. The second component explained 36.99 % of total variance with high loading of age (0.888). The communality ranged from 0.959 (HG) to 0.990 (Age) and unique factors ranged from 0.01 to 0.041 for all the six different biometric traits (Table 6). Based on Eigenvalue higher than one, only one component could be extracted based on scree plot (Fig. 3) and it explained around 93.72 % of total variance and no varimax rotation could be applied.

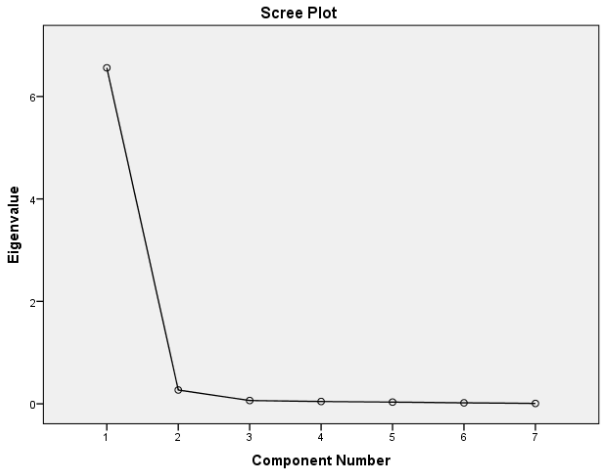

Fig. 3: Scree plot showing component number with Eigenvalues of HF crossbred male (including age as a factor)

Table 5: Total variance explained by different components in HF crossbred male (including age as a factor)

Component	Eigen value			Extraction Sums of Squared Loadings			Rotation Sums of Squared Loadings		
	Total	% Variance	Cumulative %	Total	% Variance	Cumulative %	Total	% Variance	Cumulative %
1	6.56	93.72	93.72	6.56	93.72	93.72	4.24	60.58	60.58
	0.27	3.86	97.57	0.27	3.86	97.57	2.59	37.00	97.57
3	0.06	0.92	98.49	0.06	0.92	98.49	0.06	0.81	98.38
4	0.05	0.64	99.13	0.05	0.64	99.13	0.04	0.62	99
5	0.03	0.49	99.62	0.03	0.49	99.62	0.03	0.50	99.5
6	0.02	0.27	99.89	0.02	0.27	99.89	0.02	0.30	99.8
7	0.01	0.11	100.00	0.01	0.11	100.00	0.02	0.20	100

Table 6: Varimax rotated component matrix (a) of different factors for biometric traits with its communalities and unique factor in HF crossbred male (including age as a factor) (n =72)

Traits	Component "a"		Communalities	Unique Factor
	1	2		
Age	0.448	0.888	0.990	0.010
BL	0.836	0.532	0.981	0.019
HW	0.865	0.487	0.985	0.015
HH	0.901	0.411	0.981	0.019
HG	0.768	0.607	0.959	0.041
CD	0.804	0.563	0.963	0.037
WH	0.737	0.653	0.970	0.030

For HF crossbred male without age factor, KMO & MSA was obtained as 0.909. The overall significance of the correlation matrix tested with Bertlett's test of sphericity for the biometric traits (chi-square = 1061.509, $p<0.01$) was significant. Out of six components, two principal components extracted which could explain 98.15% of cumulative variance (Table 7). First component accounted for 51.18 % of the variation. It was represented by significant positive high loading of BL (0.701), HG (0.787), CD (0.765) and WH (0.815). The second component explained 46.98% of total variance with high loading of HW (0.775) and HH (0.823). The communality ranged from 0.968 (HG) to 0.996 (HH) and unique factors ranged from 0.004 to 0.032 for all six different biometric traits (Table 8). Based on Eigenvalue higher than one, only one component could be extracted (Fig. 4). It explained around 96.43 % of total variance and no varimax rotation could be applied. Results of PCA suggested that the use of principal component one (PC1) provided a means of reduction in the number of biometric traits to be recorded in HF crossbred male with or without age factor.

Fig. 4: Scree plot showing component number with Eigenvalues of HF crossbred male (excluding age factor)

The measure of sampling adequacy, KMO, were 0.937, 0.925, 0.913 and 0.903 in female including and excluding age factor, and in male including and excluding age factor groups, respectively, which are higher than the sampling adequacy reported by Yakubu *et al.* (2009) (0.900-0.902) and Pundir *et al.* (2011) (0.891) in White Fulani cattle and Kankrej cows, respectively. The estimate of sampling adequacy, KMO revealed the proportion of the variance in different biometric traits caused by the underlying factors. Bertlett's tests of Sphericity obtained were 5091.552, 5515.716, in

Table 7: Total variance explained by different components in HF crossbred male (excluding age factor)

Component	Eigen value			Extraction Sums of Squared Loadings			Rotation Sums of Squared Loadings		
	Total	% Variance	Cumulative %	Total	% Variance	Cumulative %	Total	% Variance	Cumulative %
1	5.79	96.43	96.43	5.79	96.43	96.43	3.07	51.18	51.18
2	0.10	01.72	98.15	0.10	01.72	98.15	2.82	46.98	98.15
3	0.05	00.75	98.90	0.05	00.75	98.90	0.06	0.7	98.85
4	0.04	00.62	99.53	0.04	00.62	99.53	0.03	0.5	99.35
5	0.02	00.33	99.86	0.02	00.33	99.86	0.01	0.35	99.70
6	0.01	00.15	100.00	0.01	00.15	100.00	0.01	0.3	100

Table 8: Varimax rotated component matrix (a) of different factors for biometric traits with its communalities and unique factor in HF crossbred male (excluding age factor)

Traits	Component "a"		Communalities	Unique Factor
	1	2		
BL	0.701	0.700	0.982	0.018
HW	0.626	0.775	0.993	0.007
HH	0.564	0.823	0.996	0.004
HG	0.787	0.591	0.968	0.032
CD	0.765	0.621	0.971	0.029
WH	0.815	0.561	0.979	0.021

female including and excluding age factor, and 1185.723 and 1061.509 in male including and excluding age factor groups, respectively.

In comparison to two factors extracted, which explained 95.88 to 98.15% cumulative variance in female and male groups, Pundir *et al.* (2011) extracted three factors from 18 different biometric traits in Kankrej cows which accounted for 66.02 % of total variation, Sadek *et al*. (2006) extracted three factors for Arabian mares and stallions separately by studying 14 different traits and these explained 66% and 67% of total variation. Salako (2006) extracted two factors from 10 different biometric traits in Uda sheep which accounted for 75% of total variation.

In the present study, the first factor accounted for 69.20%, 49.35 % of the variation in female including and excluding age factor, and 60.58% and 51.17% of the variation in male including and excluding age factor groups, respectively. Pundir *et al.* (2011) reported in Kankrej cows that the first factor explained 38.89% of total variation. The first factor explaining maximum/highest variation was in accordance with the Hammack and Shrode (1986), Salako (2006), Sadek *et al*. (2006), Karacaoeren and Kadarmideen (2008), Yakubu *et al*. (2010) and Pundir *et al.* (2011). In present study, second factor accounted for 26.68 %, 47.54 %, 36.99% and 46.97 % of the variation in female including and excluding age factor, and in male including and excluding age factor groups, respectively. Yakuba *et al.* (2010) reported that the second factor explained 6.38% and 7.68% of total variation, while Salako (2006) reported that the second factor explained 11.03% of total variation in Uda sheep and Sadek *et al.* (2006) observed it as 15% and 17% of total variation in Arabian mares and stallions, respectively.

Approximate range of communality reported in Kankrej cows by Pundir *et al*. (2011) was 0.372 to 0.613. Sadek *et al*. (2006) reported approximate range of communality as 0.42 to 0.87 and 0.32 to 0.83 in Arabian mares and stallions, respectively, which were lower than present findings. In nearly accordance to the present study, Yakubu *et al.* (2010) estimated approximate communality ranging from 0.79 to 0.93 in goats.

Conclusions

Principal component analysis of morphometric traits showed that most of variation explained by PC1 in female and male groups, irrespective of age factor. Commonalities were higher which showed that all the variables were important, but PC1 had high values for BL, HW, HH, HG, CD and WH for female and male group, including age as a factor, while BL, HG, CD and WH for female and male group, excluding age factor. Shortly, if a farmer is unaware about the age of his animal still he can find accurate result by using PC1 components. This indicated that biometric traits are very important for selection of genetically elite animals. Biometric traits can be used to estimate the body weight in the field conditions, where weighing balance is not usually available.

Acknowledgement

We are thankful to Dr. S.V.Shah Professor and Head, LRS, AAU, Anand for allowing us to access the farm for data collection. We are also thankful to Dr. Sanjay B. Patel, Manager, ARDA, Amul dairy, Anand for allowing us to collect data from their Ode and Sarsa farms. We thank & acknowledge Dr. D. J. Parmar, Associate Professor, Department of Agricultural Statistics, B.A. College of Agriculture, Anand for providing the guidance for data analysis.

References

Fumio, M., Hideaki, N., & Toyokazu, F. (1982). Application of principal component analysis for defining size and shape in Japanese black cattle sires. *The Science reports of Faculty of Agriculture, Kobe University, 15,* 169-176.

Hammock, S.P., & Shrode, R.P. (1986). Calfhood weights, body measurements and measures of fatness versus criteria of overall size and shape for predicting yearling performance in beef cattle. *Journal of Animal Science*, *63*, 447-452.

Hotelling, H. (1933). The most predictable criterion. *Journal of Education Psychology, 26*, 139-142.

Johnson, R.A., & Wichern, D.W. (2007). *Applied Multivariate Statistical Analysis*. 6th ed., Upper Saddle River, NJ: Prentice-Hall

Kaiser, H.F. (1958). The varimax criterion for analytic rotation in factor analysis. *Psychometrika*, *23*(3), 187-200.

Karacaroen, B., & Kadarmindeen, H.N. (2008). Principal component and clustering analysis of functional traits in Swiss dairy cattle. *Turkish Journal of Veterinary & Animal Sciences, 32,* 163-171.

Manly, B.F. (1994). 10 ecological statistics. *Handbook of Statistics*, *12*, 307-376.

Pundir, R.K., Singh, P.K., Singh, K.P., & Dangi, P.S. (2011). Factor analysis of biometric traits of Kankrej cows to explain body conformation. *Asian-Australasian Journal of Animal Sciences*, *24*(4), 449-456.

Sadek, M.H., Al-Aboud, & Ashmawy, A.A. (2006). Factor analysis of body measurements in Arabian horses. *Journal of Animal Breeding and Genetics, 123*, 369-377.

Salako, A.E. (2006). Principal component factor analysis of the morphostructure of immature Uda sheep. *International Journal of Morphology*, *24*(4), 571-574.

Schwabe, A.E., & Hall, S.J.G. (1989). Dystociain nine British breeds of cattle and its relationship to the dimensions of the damand calf. *Veterinary Record, 125,* 636-639.

Simon, D. L., & Buchenauer, D. (1993). *Genetic diversity of European livestock breeds*. Wageningen Pers.

Yakubu, A., Ogah, D.M., & Idahor, K.O. (2009). Principal component analysis of the morphostructural indices of White Fulani cattle. *Trakia Journal of Science*, *7*(2), 67-73.

Yakubu, A., Salako, A.E., Imumorin, I.G., Ige, A.O., & Kinyemi, M.O.A. (2010). Discriminate analysis of morphometric differentiation in the West African Dwarf and Red Sokoto goats. *South African Journal of Animal Science*, *40*(4), 381-387.