PRINCIPAL COMPONENT ANALYSIS OF MORPHOMETRIC TRAITS IN KASHMIR MERINO SHEEP

ABSTRACT

Aims: To study the principal component analysis (PCA) of morphometric traits in Kashmir Merino sheep

Place and Duration of Study:SBF, KralaPathri, Kashmir, India and 2019

Methodology: This study was preformed to evaluate the morphometric traits of 518 Kashmir Merino sheep in Jammu & Kashmir under a multivariate approach. The body measurements included in the study were face length (FL), ear length (EL), ear width (EW), body length (BL), body height (BH), chest girth (CG), paunch girth (PG), tail length (TL), horn length (HL) and body weight (BW). Principal component analysis (PCA) was performed to define body shape upon the correlation matrix of the ten body measurements.

Results: Principal component analysis with varimax rotation method was applied and extracted four principal component with a total variation of 64.29%. The first principal component accounted for 28.28% of the total variance and was interpreted as a measure of CG, PG, BL and BH. The second factor which explained 12.56% of the generalized variance tended to describe TL and HL, while the third factor explained 11.88% of total variance which showed high loadings for BW. The fourth factor was influenced by FL and EW which explained 11.56 of total variance.

Conclusion: Therefore, these findings indicated that the application of PCA in breeding programs can contribute to the genetic improvement of Kashmir Merino sheep, fostering better growth performance, increased body size and enhanced productivity.

Key words:*Kashmir Merino, Morphometric traits, Principal component analysis and breeding programs*

1. INTRODUCTION

Linear body measurements, in addition to other traits, are used to distinguish various livestock breeds, helping to identify their origin, relationship, and individuals' shape and size characteristics**.** For predicting body conformation and market weight of farm animals, several linear body measurements have been found to be crucial.Body weights at different stages of life cycleare very important traits for judging the performance and adaptability of a genotype to the environment [14]. In livestock breeding programs, animals are selected based on breeding objectives that consider multiple correlated traits.Certain traits may be redundant in genetic evaluations due to their strong correlation with each other. Additionally, analyzing these correlated traits simultaneously does not enhance accuracy but instead demands large number of observations and analysis efforts[12].Kashmir Merino is an important genetic resource of Kashmir developed through crossbreeding. This sheep population is highly variable with respect to morphological structure and production traits (Rather et al. 2019).Kashmir Merino is a crossbred strain produced by crossing Gaddi, Bhakarwal and Poonchi sheep with 50% to 75% exotic inheritance from Rambouillet and Merino sheep [19]. Kashmir Merino was developed in the state of Jammu and Kashmir through crossbreeding between 1942 and 1964, with the goal of enhancing growth and adaptability. The breed was developed through crossbreeding to enhance the genetic potential of native sheep for wool traits, aiming to meet the increasing demand for high-quality apparel wool [15].Australian Merino Rams were mated with local Kashmir valley ewes, and F1 ewes were bred to American Delain Rams. F2 were bred amongst themselves following appropriate selection based on body weight and wool quality. The F2 generation kept mating until a breed known as the "Kashmir Merino" with consistent and uniform characteristics developed [1]. PCA is a widely used multivariate technique. It is a mathematical procedure that employs orthogonal transformations on the covariance matrix to reduce a set of correlated variables to a smaller set of uncorrelated variables, known as principal components, while retaining most of the original information from the covariance matrix [5]. The components are arranged in such a way that the first few capture the majority of the variation present in the original variables [7]. From the perspective of animal genetics, principal components simultaneously evaluate a set of attributes that can be utilized for selection purposes [13]. Therefore, PCA is a multivariate ordination technique used to find a linear combination of data sets that define the maximum variance [6] and transforms variables in a multivariate dataset into new, uncorrelated variables, which account for decreasing proportions of the total variance of the original variables.

2. MATERIAL AND METHOD

2.1 Description of Data: Data pertaining to morphological traits of 518 adult Kashmir Merino sheep of both sexes were collected and analyzed. The traits included in the study wereface length (FL), ear length (EL), ear width (EW), body length (BL), body height (BH), chest girth (CG), paunch girth (PG), tail length (TL), horn length (HL) and body weight (BW).

2.2 Statistical Analysis: Phenotypic correlations (r) were determined using Pearson's correlation to evaluate the degree of association among the morphological measurements. The Kaiser-Meyer-Olkin (KMO) [8] test was used to assess data adequacy, while Bartlett's Test of Sphericity [2] was applied to confirm the accuracy of the factor analysis. Initially, a dataset comprising 518 animals and ten traits underwent Bartlett's test to determine its suitability for factorization, as suggested by Maxwell. The dataset's validity was confirmed at a 1% significance level using the KMO test for sample adequacy. The number of factors was identified using the Kaiser rule criterion, which retains only factors with eigen values greater than 1. The appropriateness of the common factor model was assessed using Kaiser's Measure of Sampling Adequacy (MSA), with values below 0.5 considered undesirable. Traits showing high correlations were then analyzed using multivariate principal component analysis. The percentage of overall variation and eigen values were calculated, along with factor patterns, eigenvectors, and variable loadings. The suitability of PCA was verified using SPSS version 24 [18]. PCA is a technique, as outlined by [3], that transforms variables in a multivariate dataset (X1, X2, ..., Xn) into a set of uncorrelated variables (Y1, Y2, ..., Yn). These new variables account for a reduced proportion of the total variance compared to the original variables, which were specified as follows:

$$Y_1 = a_{11}X_1+a_{12}X_2+\ldots+ a_{1n}X_n.$$

$$Y_2 = a_{21}X_1+a_{22}X_2+\ldots+ a_{2n}X_n.$$

$$Y_n = an_1X_1+an_2X_2+\ldots+ a_nX_n.$$

The original variables X1, X2,..., Xn include larger amounts of variation than the principal components Y1, Y2,..., Yn. An orthogonal rotation that maximizes the variance was used to linearly modify the factor pattern matrix in order to make it easier to understand and interpret these primary components.

3. RESULT AND DISCUSSION

Principal component analysis was applied to ten different morphological traits in Kashmir Merino sheep. The Kaiser-Meyer-Olkin method yielded a measurement of sampling adequacy of 0.746 (Table 1) whereby eigen values more than 1 were considered. This value assesses the adequacy of the data for each factor in providing reliable results for PCA. The measure of sampling adequacy below 0.5 is considered to be inadequate [9]. The varimax rotation method was employed to maximize the sum of loading squares [4]. Bartlett's test of sphericity was applied to assess the significance of the correlation matrix. The resulting chi-square value was highly significant (P<0.01), measuring 1174.437 (Table 1). The sum of squares loadings were extracted by PCA, variation explained by each component (Table 2) and eigen values are given in scree plot (Fig 1). Four, out of the ten components were selected using the Kaiser Rule Criterion [6] to determine the significant number of components. The cumulative variance of 64.29% (Table 2) was accounted by four principal components (PC1, PC2, PC3 and PC4), each with eigen values more than 1. The variance of each trait was as per explained by PCA [11]. The component plot in rotated space depicts the distribution of the ten components, as illustrated in Figure 2. The plot is a result of varimax rotation, a method that enhances interpretability by maximizing the variance explained by each component. This technique helps to understand which traits are most closely associated with each component, making the overall analysis more comprehensible.

Table 1. KMO and Bartlett's Test

Kaiser-Meyer-Olkin Measure of Sampling Adequacy.		0.746
Bartlett's Test of Sphericity	Approx. Chi-Square	1174.437
	Df	45
	Sig.	0

In the present study, the first principal component (PC1) explained for 28.28% of the total variance (Table 2). PC1 represented by a very high component loading for CG and PG in Kashmir Merino sheep. The second principal component (PC2) explained for 12.56% of total variance. PC2 showed high loadings on TL and HL. The third principal component (PC3) explained 11.88% of total variance which described high loads on BW. The fourth principal component (PC4) explained for 11.56% of total variance which described high loads on FL. Table 4 represents the coefficient of principal component analysis of rotated component matrix. The different weights were assigned by PC1, PC2, PC3 and PC4 showing different component loadings in Kashmir Merino sheep.

Table 3 represents the communality values ranged from 0.518 (EW) to 0.835(CG) across all economic traits. The trait like EW had lower communality indicating that this trait is less effective in explaining the performance in Gurez sheep and traits including CG and PG showed high communalities indicating that these traits will be effective in breeding programme for selection of Kashmir Merino sheep.

Table 2. Total variance explained in different morphometric traits

Component	Initial Eigen values			Extraction Sums of Squared Loadings			Rotation Sums of Squared Loadings		
	Total	% of Variance	Cumulative %	Total	% of Variance	Cumulative %	Total	% of Variance	Cumulative %
1	2.905	29.047	29.047	2.905	29.047	29.047	2.828	28.280	28.280
2	1.267	12.668	41.715	1.267	12.668	41.715	1.257	12.567	40.847
3	1.237	12.367	54.082	1.237	12.367	54.082	1.189	11.888	52.734
4	1.021	10.212	64.295	1.021	10.212	64.295	1.156	11.560	64.295
5	.928	9.282	73.577						
6	.783	7.834	81.411						
7	.715	7.146	88.557						
8	.521	5.206	93.763						
9	.443	4.425	98.188						
10	.181	1.812	100.000						

Similar findings were reported by [20] which revealed two components accounted for a substantial proportion of the total variance in both Balami and Uda sheep. Specifically, the components explained 66.91% of the total variance for Balami sheep and 57.43% for Uda sheep. The KMO measures of sampling adequacy were 0.923 for Balami sheep and 0.932 for Uda sheep, indicating a very high adequacy of the samples for factor analysis. Communalities after extraction ranged between 0.441 and 0.903 for both genotypic groups, suggesting that a significant portion of the variances is shared among the variables. This supports the suitability of using PCA to classify these sheep based on the analyzed traits.In a study of Kail sheep [10] reported PCA in morphological trait which extracted two components and explaining total variance of 67 % and 60%, respectively. PC1 is positively associated with live body weight and heart girth, indicating that these traits are significant contributors to this component. Meanwhile, PC2 accounts for

28% of the total variance and is positively correlated with tail length and head length. [16] conducted PCA on the morphological structure of Uda sheep, identifying two principal components that together explained 67.6% and 11.03% of the generalized variance in body measurements, respectively. The first component included measurements closely linked to bone growth, such as forelimb length, tail length, facial length, rump height, withers height, and body length. The second component comprised dimensions that are less associated with bone growth, including rump width, shoulder width, and rump length.The Kaiser-Meyer-Olkin (KMO) measure of sampling adequacy was 0.85, indicating a high proportion of the variance in the body measurements attributable to underlying factors. The communalities for the variables ranged from 0.61 to 0.93, demonstrating substantial shared variance among the body measurements. PCA in morphological traits of Rampur-Bushair sheep [17] which described three and four factors in young and adult sheep respectively, which accounted for 57% and 61% of variation. The extracted principal component significantly contributed to explaining the overall body conformation. In an earlier study of Zulu sheep [11] reported PCA explaining a total variance of 66.85% in young sheep and four components in adult sheep, accounting for 62.13% of the total variance. In both cases, PC1 exhibited high loadings for variables associated with body size, whereas PC2 was strongly correlated with traits reflecting body shape.

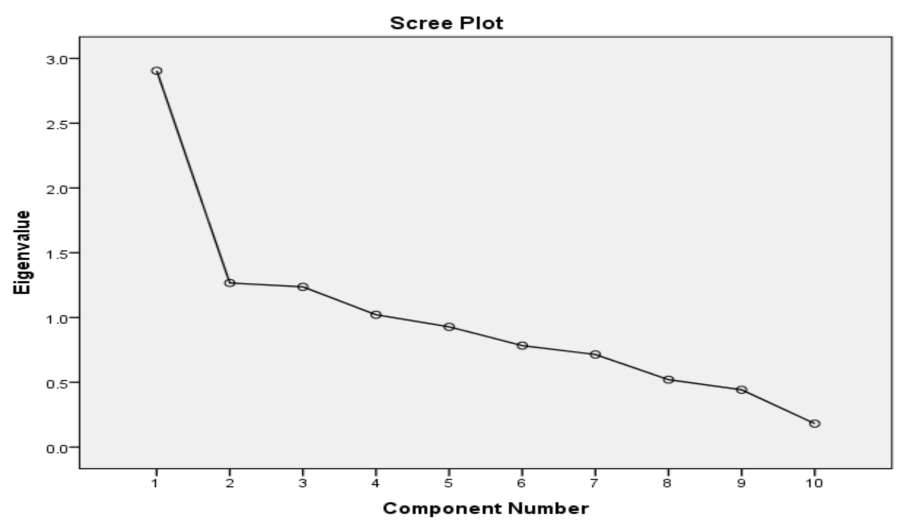

Fig 1. Scree Plot

Table 3. Communalities of different morphometric traits

	Initial	Extraction
FL	1.000	.578
EL	1.000	.540
EW	1.000	.518
BL	1.000	.596
BH	1.000	.608
CG	1.000	.835
PG	1.000	.764
TL	1.000	.672
HL	1.000	.650
BW	1.000	.669

Table 4. Rotated Component Matrix of different factors for morphometric traits

Trait	Component			
	1	2	3	4
FL	0.095	0.012	0.181	0.732
EL	0.198	-0.125	-0.646	-0.26
EW	0.077	-0.052	-0.103	0.706
BL	0.77	0.006	-0.034	0.053
BH	0.763	0.122	0.080	0.075
CG	0.906	0.022	-0.086	0.078
PG	0.871	0.032	-0.052	0.040
TL	0.1	0.788	-0.186	0.080

HL	0.032	0.77	0.201	-0.127
BW	0.101	-0.089	0.797	-0.123

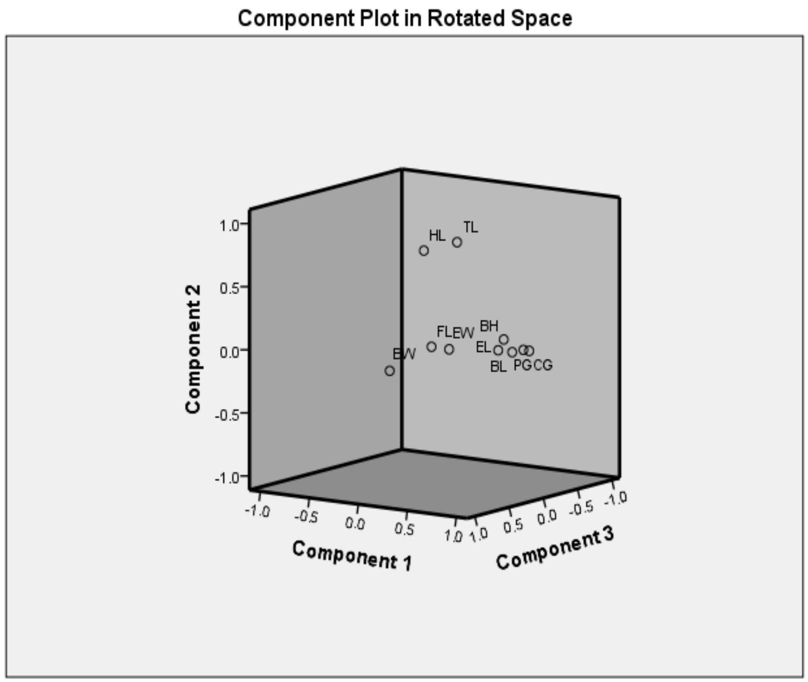

Fig 2. Rotated Component Plot

4. CONCLUSION

The principal component analysis demonstrated that a few principal components could explain a substantial proportion of the total variability in the morphometric traits. This indicates that a reduced set of traits can serve as effective predictors for the overall body conformation of the sheep. Traits like chest girth and body length emerged as critical contributors to the first principal component, highlighting their importance in assessing the overall size and productivity potential of Kashmir Merino sheep. Overall, the study of morphometric traits in Kashmir Merino sheep provides a robust foundation for future breeding programs aimed at improving their genetic potential. By focusing on key morphometric indicators, breeders can achieve significant gains in body conformation and overall productivity, contributing to the sustainable development of the sheep population in Kashmir.

REFERENCES

1. Anonymous. Directorate of Sheep Husbandry Kashmir Division, Government of Jammu and Kashmir. 2004; (assessed 2nd January 2018). http:// jksheephusbandrykashmir.net/sheep.html.
2. Bartlett MS. Tests of significance in factor analysis, The British Journal of Psychology. 1950; 77-85.
3. Everitt BS, Landau S, Leese M. 2001. Cluster analysis. 4th edn., Arnold Publisher, London.
4. Fernandez G. Data Mining using SAS Application. USA: Chapman and Hall, CRC press; 2002.
5. Hair JF, Black WC, Babin BJ, Anderson RE. Multivariate Data Analysis. 2009; 7th ed. Prentice Hall, Upper Saddle River.
6. Johnson RA, Wichern DW. Applied Multivariate Statistical Analysis. USA: Prentice-hall Inc. 1982.
7. Jolliffe I. Principal Component Analysis. 2002; 2nd ed. Springer.
8. Kaiser H F. The varimax criterion for analytic rotation in factor analysis. Psychometrika. 1958; 23: 187–200.
9. Khargharia G, Kadirvel G, Kumar S, Doley S, Bharti PK, Das M. Principal component analysis of morphological traits of Assam Hill goat in Eastern Himalayan India. JAPS: Journal of Animal & Plant Sciences. 2015; 25(5):1251-1258.
10. Kuthu ZH, Ahmed Z, Nawaz M, Bilal G, Zubair M, Ahmed I, Ikram M (2022). Principal Component Analysis of Kail sheep based on body measurements. Pakistan Journal of Science, 2022; 74(4):377.
11. Mavule BS, Muchenje V, Bezuidenhout CC, Kunene NW. Morphological structure of Zulu sheep based on principal component analysis of body measurements. Small Ruminant Research. 2013; 111: 23-30.
12. Panda S, Gaur GK, Sahoo NR, Bharti PK, Kar J. Principal component analysis of morphometric and growth traits in crossbred piglets. Indian Journal of Animal Sciences. 2020; 90(8):1168-1171.
13. Pinto L FB, Packer IU, DeMelo CMR, Ledur MC, Coutinho LL. Principal components analysis applied to performance and carcass traits in the chicken. Animal Research. 2006; 55:419-425.
14. Rather MA, Bashir I, Hamdani A, Khan NN, Ahangar SA, Nazki M. Prediction of body weight from linear body measurements in kashmir merino sheep. Adv. Anim. Vet. Sci. 2021; 9(2):189-193.
15. Rather MA, Shanaz S, Ganai NA, Baba MA, Hamadani A, Ahmad M S, Ahmad T. Genetic, phenotypic and environmental trends for production and reproduction traits in Kashmir Merino Sheep. International Journal of Livestock Research. 2019; 9(12):73-78.

16. Salako AE. Application of morphological indices in the assessment of type and Function in sheep. International Journal of Morphology. 2006; 24(8): 13.
17. Sankhyan V, Thakur YP, Katoch S, Dogra PK, Thakur R. Morphological structuring using principal component analysis of Rampur-Bushair sheep under transhumance production in western Himalayan region, India. Indian Journal of Animal Research. 2018; 52(6):917-922.
18. SPSS. 2008. Computer Software 17.0 SPSS inc., Chicago, Illinois-60606, USA.
19. Tomar, S. S. Textbook of Animal Breeding. Kalyani Publishers, Ludhiana. 2004.
20. Yunusa AJ, Salako AE, Oladejo OA. Principal component analysis of the morphostructure of Uda and Balami sheep of Nigeria. International Research Journal of Agriculture Sciences. 2013; 1(3): 45-51.