Indian J. Anim. Prod. Mgmt. Vol. 29 (3-4) 46-49 (2013)

PRINCIPAL COMPONENT REGRESSION ANALYSIS (PCRS)- AN APPROACH FOR LIFETIME MILK YIELD PREDICTION IN MURRAH BUFFALOES

T. A. KHAN, A. K. S. TOMAR, TRIVENI DUTT AND BHARAT BHUSHAN
Indian Veterinary Research Institute, Izatnagar-243 122

Received : 21.10.2013 ***Accepted : 20.01.2014***

ABSTRACT

Principal Component Regression Analysis (PCRA) has been carried out to formulate lifetime milk yield prediction model with principal components as predictors. Principal compoents (PCs) were based on initially expressed reproductive traits and part lactation records from 10 years data (1-4-1999 to 31-3-2009) of Murrah buffaloes, of Cattle of Buffalo farm of IVRI. Models for lifetime milk yield LTMY4 (lifetime milk yield as total milk yield up to four lactations) has been evolved with retained four PCs (explained 86.67% variation of original data). Seven types of model have also been fitted to have best model for LTLMY4 with first PC_1 as predictor. The model for LTLMY4 could explained 43.30% variation in the estimated values with adjusted R^2=39.52%. Curve estimation analysis shows appropriateness of the Logarithmic function (adjusted R^2-value- 43.05%) followed by Cob-Douglas function (42.89%).

Key words: Lifetime milk production, Murrah buffalo, PCRA.

Principal component analysis (PCA) is a mathematical procedure that transforms a number of correlated variables into a smaller number of uncorrelated variables called principal components[3,5,8,9]. The objective of principal component analysis is to reduced the dimensionality (number of variable) of the data set but retain most of the original variability in the data. [2] Investigated the relationship between body condition scores and milk yield of high yielding Holstein dairy cows using multiple linear regression and principal component analysis. Principal component regression analysis (PCRA) can be carried out using principal component as regressors in place of original variables. PCA was also found advantageous over multiple regression analysis in livestock management research[1,10]. Many worker studied lifetime mik prediction based in early expressed growth, production and reproduction traits in dairy cattle (both in native and crossbred)[1,6,7,11].

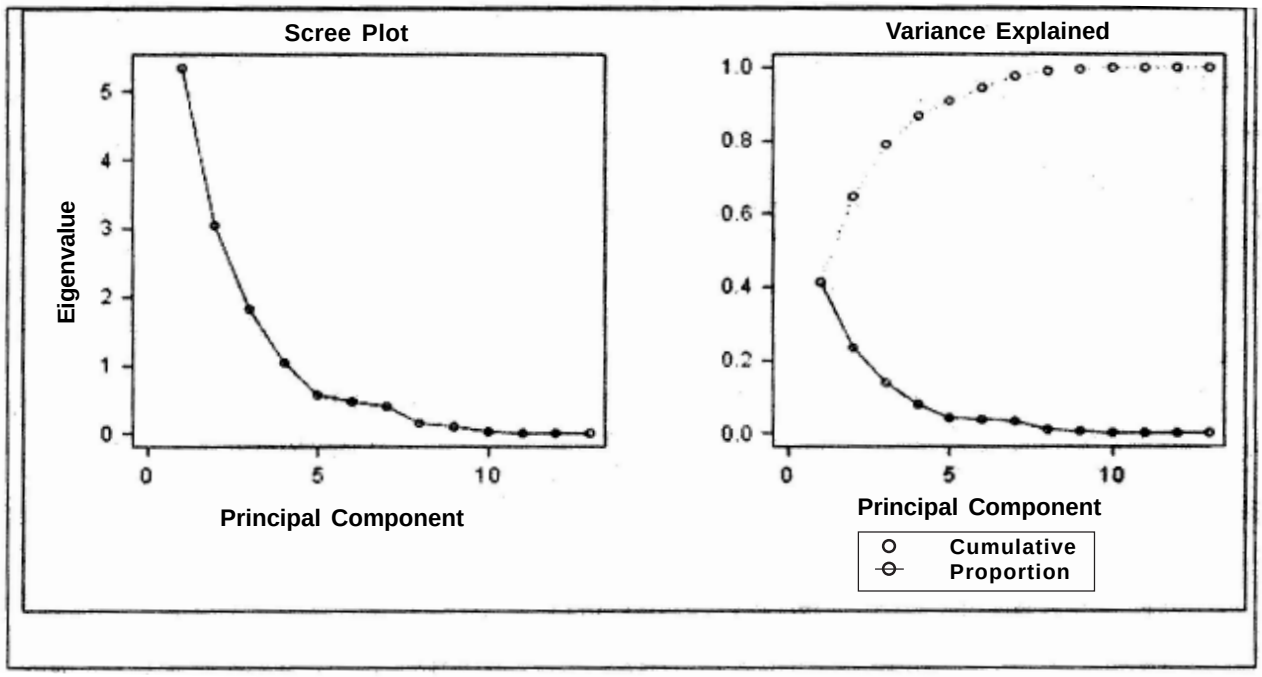
MATERIALS AND METHODS

Principal components were formulated on data on initially expressed reproductive traits (AFC-age at first calving, FLL-First lactation length, FCI-First Calving interval, FSP-First dry period) and part lactation records of 100, 170 and 240 days of first lactation and second lactation and their respective total milk yields from 10 years data (1-4-1999 to 31-3-2009) of Murrah buffaloes, at Cattle & Buffalo farm of IVRI. These components were sued as predictors for LTMY4 (lifetime milk yield as total milk yield up to four lactations) using

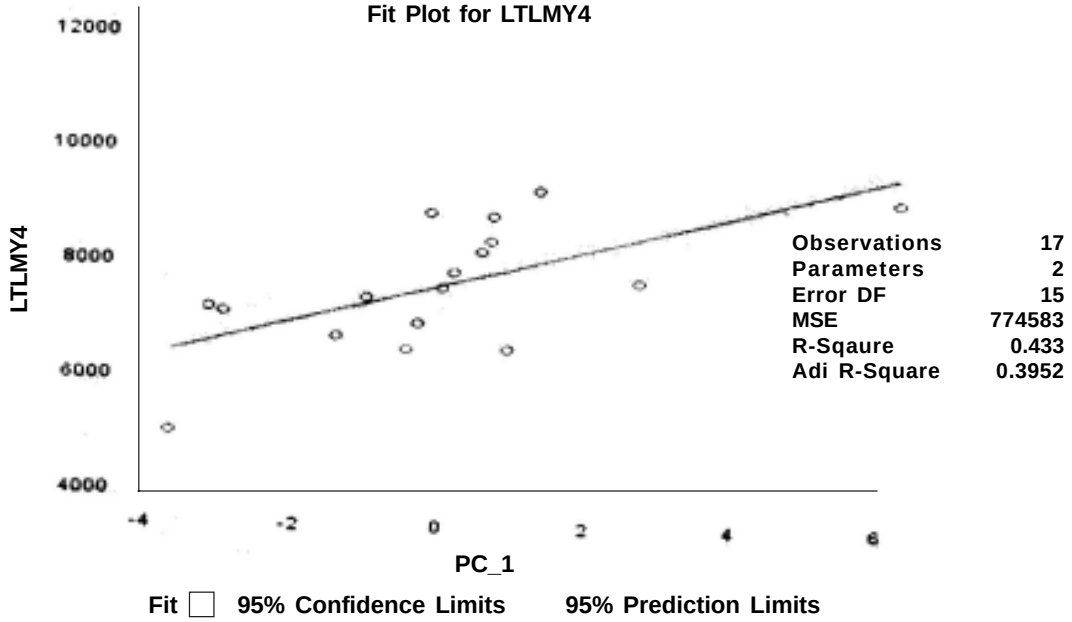
principal component regression analysis (PRA). Seven types of model were also fitted to know the best fitted model. *Princomp* procedure of SAS 9.2 was used and 13 PCs were formulated12. Based on eigen values and scree diagram (elbow method) the PCs were retained to be used as predictors for lifetime milk yield.

RESULTS AND DISCUSSION

The Principal component analysis (scree plots) could retain first four PCs explaining 86.6% variation of the original variable (Fig.1) The Stepwise regression analysis LTMY4 could retain PC_1 only.



The Model for LTMY4 (kg)=7633.418 + 321.788 PC_1, explained 43.30% variation in the estimated values with adjusted R2=39.52%. The fit-plot has shown the appropriatness of the model (Fig.2).



Curve estimation analysis with PC_1 as lone predictor; have been taken for prediction of LTMY4. Seven types of models (viz; linear, quadratic, cubic, logarithmic, exponential and Cob-Douglas) have been evolved and based on significance of regression and maximum adjusted R^2-values best models was decided. In all the evolved models the regressions of LTLMY4 with PC_1 were highly significant ($p<0.01$). The results shows that the Logarithmic function expresses maximum adjusted R^2-value (43.05% followed by Cob-Douglas function (42.89%) (Table1).

Table 1: Showing significance of co-efficients of the models for LTLMY4) and respective Adjusted R^2-values.

Model	b_0	b_1	b_2	b_3	R^2	Adj. R^2
Linear $Y\ (LTLMY4)=b_0+b_1\ X(=PC\_1)$	7633.417** (213.456)	321.788** (95.076)	-	-	43.30	39.52
Quadratic $Y=b_0+b_1 X+b_2 X^2$	7757.326** (248.737)	365.990** (105.477)	-24.582^{NS} (25.216)	-	46.90	39.32
Logarithmic $Y=b0+b1\ Log_e(10+PC\_1)$	-78.308^{NS} (2141.045)	3385.147** (935.427)	-	-	46.61	43.05
Cubic $Y=b_0+b_1 X+b_2 X^2+b_3X^3$	7801.447** (276.448)	271.815^{NS} (246.221)	-43.574^{NS} (51.578)	5.693^{NS} (13.357)	47.64	35.55
Exponential $Log_eY=Log_e(b0)+b1\ X$	7549.224** (1.030)	0.044** (0.013)	-	-	42.03	38.17
Cob-Douglas $Y=b_0(10+PC\_1)b_1$	2608.107** (1.344)	0.467** (0.129)	-	-	46.46	42.89

Figures in parenthesis are standard error of coefficient.
** & *=significant (<0.01) & (p<0.05)
The Fit Plot for Logarithmic as well as Cob-Douglas Function shows the fitness of the model with the observed LTLMY4 (Fig. 3 & 4).

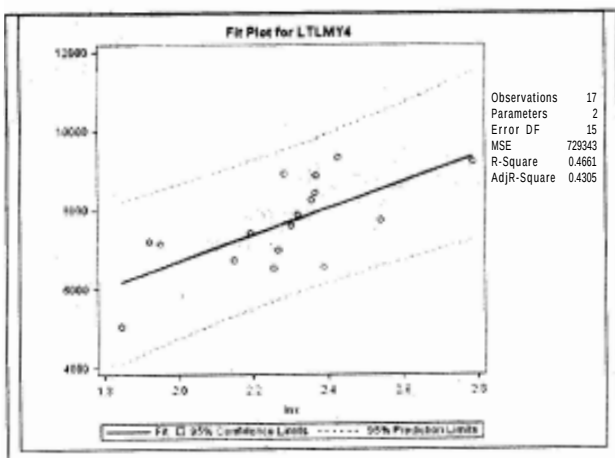


Fig.3: Showing appropriateness of Logarithmic Function.

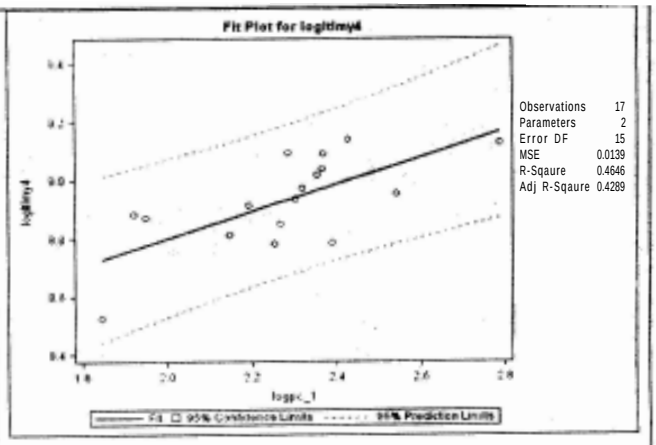


Fig.4: Showing appropriateness of Cob-Douglas Function with the observed data.

Since Murrah buffalo is one important milk contributor to the national milk production, its selection based on initially expressed traits is very important. Principal component regression analysis is necessitated for Murrah as buffaloes are being major contributor in milk production of the country, their selection at the initital stage is important. To incorporate as possible as all the available traits, PCRA can reduced the information on large number of traits on small number of animals, into principal component, expressing the amount of variation of original data. At IVRI Farm, herd is a mixture of purchased as well home-grown one; the data on initially expressed traits was less in numbers.

The logarithmic function LTLMY4=-78.308+3385.147 loge (10+PC_1), have been found best in prediction of lifetime milk yield in murrah buffalo. The prediction model, obtained through principal component regression techniques can be used as tools for early selection of animals.

REFERENCES

1. Bhattacharya T.K. and gandhi R.S. 2005. Principal components ersus multiple regression analysis to predict lifetime production of Karan Fries cattle Indian Journal of Animal Sciences 75 (11) : 13 17-1320, November, 2005.
2. Chapman K.W., Lawless H.T., and Boor K.J. 2001. Quantitive Descritive Analysis and Principal Component Analysis for Sensory Characterization of Ultrapasterized Milk J. Dairy Sci. 84:12-20.
3. Hotelling H. 1933. Analysis of a Caomplex of Statistical variables into Principal Components. Journal of Educational Psychology 24:41-41, 498-520.
4. Khan T.A. Tomar A.K.S. and Dutt triveni 2012. Prediction of Lifetime milk production in synthetic crossbred cattle strain Vrindavani of North India Indian Journal of Animal Sciences, 82:1367-1371.
5. Lukbisi, F.B., W.B. Muhuyi, J.M.K. Muia, S.N. Ole Sinkeet and W.F. Wekesa 2008a. Statistical use and Interpretation of Principal Component Analysis in Applied research. Egerton University's 3rd Annual Reserach Week and International Conference, 16-18 September 2008.
6. Malhotra, P.K. and Singh, R.P. 1980. Estimation of life-time production in Red Sindhi cattle using ridge-trace criterion. Indian Journal of Animal Sciences 50(3):215-218.
7. Puri T.R. and Sharma KNS 1965. prdiction of Lifetime Production on Basis of First lactation Yield and Age at First Calving for Selection of Dairy Cattle, Journal of Dairy Science 48(4):462-467.
8. Pearson, K. 1901. On Lines and Planes of Closet fit to a system of points in Space. Philosophical Magazine 2:557-72.
9. Rao, C.R. 1964. The use and Interpretation of Principal Component analysis in Applied Reserach, Sankhya A 26:329-358.
10. Rougoor CW, Sundaram, R, Van Arendonk JAM 2000. The relation between breeding management and 305 day milk production, determined via princiapl components regression and partial least squares. *Livestock Production Science,* 66:71-83.
11. Shinde, N.V. Mote, M.G. Khutal B.B. and Jagtap D.Z. 2010 prediction of Lifetime Milk production on the basis of lactatio traits in Phule Triveni Crossbred Cattle. Indian Journal of Animal Sciences 80 (10):968-88.
12. SAS Institute Inc. 2009. SAS/STAT9.2 User's Guide, Second Edition. Cary, NC:SAS Institure Inc.

★ ★ ★