

Bartlett Factor Scores in Multiple Linear Regression Equation as a Tool for Estimating Economic Traits in Broilers

Oluwatosin M. A. Jesuyon.

Abstract— In order to propose a simpler tool that eliminates the age-long problems associated with the traditional index method for selection of multiple traits in broilers, the Bartlett factor regression equation is being proposed as an alternative selection tool. 100 day-old chicks each of Arbor Acres (AA) and Annak (AN) broiler strains were obtained from two rival hatcheries in Ibadan Nigeria. These were raised on deep litter system in a 56-day feeding trial in the University of Ibadan Teaching and Research Farm, located in South-west Tropical Nigeria. The body weight and body dimensions were measured and recorded during the trial period. Eight (8) zoometric measurements namely Live weight (g), Abdominal Circumference, Abdominal length, Breast width, leg length, Height, Wing length and Thigh circumference (all in cm) were recorded randomly from 20 birds within strain, at a fixed time on the first day of the new week respectively with a 5-kg capacity Camry scale. These records were analyzed and compared using completely randomized design (CRD) of SPSS analytical software, with the means procedure, Factor Scores (FS) in stepwise Multiple Linear Regression (MLR) procedure for initial live weight equations. Bartlett Factor Score (BFS) analysis extracted 2 factors for each strain, termed Body-length and Thigh-meatiness Factors for AA, and; Breast Size and Height Factors for AN. These derived orthogonal factors assisted in deducing and comparing trait-combinations that best describe body conformation and Meatiness in experimental broilers. BFS procedure yielded different body conformational traits for the two strains, thus indicating the different economic traits and advantages of strains. These Factors could be useful as selection criteria for improving desired economic traits. The final Bartlett Factor Regression equations for prediction of body weight were highly significant with $P < 0.0001$, R^2 of 0.92 and above, VIF of 1.00, and DW of 1.90 and 1.47 for Arbor Acres and Annak respectively. These FSR equations could be used as a simple and potent tool for selection during poultry flock improvement, it could also be used to estimate selection index of flocks in order to discriminate between strains, and evaluate consumer preference traits in broilers.

Keywords—Live body weight, Linear and body measurements, Alternative selection tool, Bartlett factor regression model, Consumer preference traits.

I. INTRODUCTION

POULTRY industry is one of the fastest growing segments in the agricultural sector and it undoubtedly plays an important role in the economy [1] of Tropical countries. Current trend in the industry is that the market live weight of broiler determines its commercial value in Tropical locations. Morphometric measurements are reported useful in contrasting size and shape of animals [2], [3], [4], but correlations among these body measurements may cause lack of orthogonality or collinearity among explanatory variables due to the interrelationship among traits. Principal component analysis (PCA) is a multivariate technique that reduces variables and breaks multicollinearity among inter-related traits in a data set. Resulting factor scores (FS) are highly correlated with specific variables and are unbiased estimates of the true factor scores. They are linear combinations of the observed variables and could be utilized as independent and uncorrelated values, to regress linear equations with higher precision and predictive ability. BFS is a refined procedure, among two others, applied when both principal components and common factor extraction methods are used with Explorative Factor Analysis (EFA). The three advantages of BFS over Regression Scores and Anderson-Rubin Scores are high correlation to estimated factors (high validity), non-correlation to other orthogonal factors (univocality) and production of unbiased estimates of factor score parameters. When used in MLR equation BFS become standardized observed values of the variables constituting the data sets [5]. Multiple linear regression (MLR) is a valuable tool for estimating, predicting and forecasting purposes while Factor Score Regression (FSR) method is a three-step approach [6] built on the assumption that scores resulting from factor rotation are uncorrelated in MLR. Thus FSR could be used to identify

best combining groups of traits (Factors), among many, that best define live body conformation (size and shape) and compare conformation between individuals. It could also be useful for identifying orthogonal body traits, predicting live weight and comparing between live weight equations for selection. Values from FSR equation could be useful indices for selecting individuals in broiler improvement programme. Research workers have utilized principal component analysis (PCA) to investigate the relationship among body measurements in indigenous chickens [7], [8], [9], [10], [11] and turkey [12] while two groups have furthered the use of FSR for prediction purposes in poultry [13], [14]. FSR has also been used as predictors of total carcass muscle, fat and bone, and other performance traits [15], [16], [17], [18], [19], [20]. FS has been used as a selection criterion for genetic improvement of muscle weight distribution [21]. The study aimed to identify composite body traits (principal factors) that best describe broiler chicken conformation to meet consumer preference; regress live weight equations with Factor Scores using MLR model and investigate the model for use as a simpler alternative selection tool. The experiment was conducted on the Teaching and Research farm of the University of Ibadan; Ibadan, Nigeria located on Latitude 7.4417° N, longitude 3.9000° E. The hypothesis tested was that, there was no significant difference ($p < 0.05$) between 8-week broiler weight of strains.

$$H_0: LWT_{AA} = LWT_{ANNAK}$$

$$H_1: LWT_{AA} \neq LWT_{ANNAK}$$

II. MATERIALS AND METHODS

100 day-old chicks each of Arbor Acres (AA) and Annak (AN) broilers were purchased from 2 different Hatcheries in Ibadan Nigeria. Both strains were subjected to same standard management practices prescribed for broiler chickens from day-old to 56th day. Vaccinations were administered as recommended. Generous and equal amount of formulated feed and water was offered daily to each strain. The calculated protein content of the Starter and finisher diets was 23.80 and 20.01 % while the energy content was 3025.00 and 2607.00 KCal/kg respectively. Trial was terminated on the 56th day.

Measurement of Zoometric Traits

Body weight of chicks at day-old was estimated by weighing individually with a sensitive egg scale. Weekly body weight and measurements were taken early in the morning before feeding at a fixed day and time. From the second week to the 8th week, eight (8) biometric measurements were taken of live body weight (LBW), Breast width (BRW), Wing length (WGL), Abdominal length (ADL), Abdominal Circumference (ADC), Leg length (LGL), Thigh circumference (TCF) and Height (HT). Measurements were taken from 20 birds randomly sampled from each strain, without replacement until all measurements were taken successfully. Linear measurements were made with a thin thread and the lengths were determined on a metric ruler in centimeters. Body weight was determined with a 5-kg capacity Camry dial spring scale. Reference points were as described below:

Live weight: Weight of the live bird when placed on the weighing scale

Wing length: Length from the tip of the wing to the shoulder joint.

Abdominal Circumference: Circular distance round the body of the chicken passing through the uropygeal or preening gland and the tip of the Sternum at the chest.

Abdominal length: Length from the tip of the sternum through the mid-region of the belly to the vent.

Thigh Circumference: Distance round the thigh at the thickest part near the hock joint.

Breast Width: Distance across the chest from the right wing-tip to the left wing-tip.

Height: Distance from the floor level to the mid-back line of the chicken.

Leg Length: The length of the leg from the floor level to the joint at the top of the femur (hip).

The experimental design was completely randomized design (CRD). The statistical model for the experiment was:

$$Y_{ij} = \mu + \alpha_i + \epsilon_{ij}$$

Where

Y_{ij} = Individual live body weight in i th strain and j th replicate.

μ = Overall mean

α_i = Strain (treatment) effect, where $i=1, 2$.

ϵ_{ij} = Random error component in treatment i , and replicate j , where ($i=2, j=20$)

Statistical Analytical procedure

All measurements were subjected to means procedure, PCA and Multiple linear Regression (MLR) procedures to predict body weights in the strains with standardized Bartlett Factor Scores (BFS). All measurements for each strain except live weight were subjected to the principal component analysis (PCA) procedure to extract two factors and generate scores by Bartlett's method, for use as independent variables in subsequent MLR. Factor scores obtained were then used as independent variables in MLR procedure to obtain live weight equations for each strain. SPSS version 17 [22] was used for the analysis. The model for the normal regression of body weight on other linear body parameters was:

$$Y_{ijkl} = a + b_1X_1 + b_2X_2 + \dots + b_nX_n + \epsilon_{ijkl}$$

Where

Y_{ijkl} = Response in body weight (g)

a = Constant for a strain

b = Coefficient of parameters for $i = 1$ to n

X = Body parametric measurements for $i=1$ to n .

1 to n = number of parameters extracted by the stepwise method.

ϵ_{ijkl} = Random error component

The model for the regression of live body weight on BFS was:

$$Y = \mu + BF_1X_1 + BF_2X_2 + \dots + BF_nX_n + \epsilon$$

Where

Y = Live weight of broiler (g)

μ = Strain Constant.

BF_{1-n} = Bartlett factor coefficient obtained from the PFA analysis of Experimental data.

X_{1-n} = Correlated Traits of interest in which continuous improvement is desired as grouped into Factor

components.

ε = error term of the model

III. RESULTS

Table 1 shows results of MLR analysis on both strains using the raw data. These revealed high standard error of estimate, SEE (174.59 vs. 157.59), Durbin-Watson statistics, DW (1.95 vs 1.56), high R² (0.93 vs 0.96), and high Variance Inflation Factor, VIF (12.26 vs 13.98) larger than 10 for the equations of AA and AN strains. In order to eliminate multi-collinearity observed among predicting variables in our normal regression

equations, LBW was regressed on standardized Bartlett Factor scores in the estimation equations for both strains. The Kaiser-Meyer-Olkin (KMO) measure of sampling adequacy values for Arbor Acres and Annak strains were 0.947 and 0.934. Results of Bartlett's Test of Sphericity were 997.01 and 918.05 respectively. These implied that the data sets were suitable for factor analysis [23].

Table 1: Normal Multiple Linear Regression Equations by Stepwise Method for prediction of 8-week finisher live weight based on Body traits for Arbor Acres and Annak Broilers

Strain	Parameters	Unstd. Coeff.	Std Coeff (β)	SE	t-value	P-value	VIF
Arbor Acres	Const.	-743.210	-	104.193	-7.133	0.001	-
	LGL	83.378	0.496	18.985	4.392	0.001	12.263
	BRW	49.565	0.462	11.801	4.200	0.001	11.611
	TCF	59.887	0.300	21.339	2.806	0.007	10.957
	WGL	-44.314	-0.282	17.043	-2.600	0.012	11.265
Annak	Const.	-872.312	-	80.185	-10.879	0.001	-
	ADC	32.023	0.552	5.038	6.356	0.001	13.861
	BRW	47.772	0.458	9.097	5.252	0.001	13.978
	ADL	-18.682	-0.186	6.424	-2.908	0.005	7.482
	LGL	26.966	0.163	11.417	2.362	0.021	8.733

Note: Model Adj. R² = 0.928 (0.962); SEE = 174.585 (108.877); DW = 1.949 (1.966); Model sig. P = 0.01 (0.001); BRW = Breast Width; LGL = Leg Length; WGL = Wing Length; ADC = Abdominal Circumference; ADL = Abdominal Length; TCF = Thigh Circumference; HT = Height. Values in bracket are for Annak Broilers.

Table 2 presents the component traits and predicted BFS matrix of AA and AN strains. This shows the two factors extracted and associated variance predicted for each factor, for each strain. Both factors extracted for each strain were responsible for 96.3% and 95.1% of the shared variations among linear body measurements. Extracted variance for AA was equally loaded

between BF1 and BF2 (48.26 vs 48.04) while variance in AN was primarily loaded on BF1 (92.97). The scores for these two independent Bartlett factors were saved for each strain and thereafter employed as independent and orthogonal predictors to regress standardized live weight equations.

Table 2: Predicted Rotated Bartlett Factor Scores for Component Traits of Arbor Acres and Annak Broilers

STRAIN	ARBOR ACRES		ANNAK	
	BFS1	BFS2	BFS1	BFS2
BRW	-0.407	0.616	0.696	-0.622
LGL	-1.001	1.209	0.582	-0.486
WGL	0.468	-0.259	0.300	-0.133
ADC	0.276	-0.067	0.248	-0.070
ADL	1.247	-1.043	-0.173	0.449
TCF	-0.577	0.786	0.379	-0.231
HT	0.708	-0.501	-1.273	1.805
% Variance	48.264	48.039	92.966	2.109
Total Variance Predicted	96.303		95.075	
Factor Name	Body length	Meatiness	Breast Size	Height

Table 3 displays the resulting Bartlett Factor Score Regression (BFSR) equations. The procedure yielded VIF values lower than 10 and higher t-statistics values. There were little reductions in DW (1.90, 1.48 vs 1.95, 1.97) and in R² (0.92, 0.94 vs 0.93, 0.96), and corresponding increases in SEE, (180.49, 140.09 vs 174.59, 108.88) compared with normal regression equations. The selected factors had significant and positive linear relationships with live weight (P<0.0001) in the model for each strain. Resulting standardized live weight equations incorporating the composite traits (Factors) were:

$$Y = 825.5 + 347.6 F_1 + 521.6 F_2 \pm 65.03 \quad (R^2=0.92)$$

.....Arbor Acres Strain

$$Y = 759.2 + 373.6 F_1 + 472.9 F_2 \pm 54.91 \quad (R^2=0.94)$$

.....Annak Strain

where:

Y = Live body weight in grams

F₁ = Factor 1 component.

F₂ = Factor 2 component.

Table 3: Bartlett Factor Scores in Multiple Linear Regression Equations for Prediction of Live weight for Arbor Acres and Annak Broilers

Strain	Predicting Factor	Unstd. Coeff.	Std. Coeff.	SE	t-value	P-value	VIF
Arbor Acres	Constant	825.514	-	21.573	38.267	0.0001	-
	BFS1	347.554	0.533	21.728	15.995	0.0001	1.000
	BFS2	521.566	0.801	21.728	24.004	0.0001	1.000
Annak	Constant	759.166	-	18.112	41.915	0.0001	-
	BFS1	373.565	0.707	15.868	23.542	0.0001	1.000
	BFS2	472.932	0.678	20.933	22.592	0.0001	1.000

Notes: Adj. R²= 0.923 (0.938), SEE=180.490 (140.087), DW= 1.900 (1.481), Model sig. P = 0.0001 (0.0001). BFS = Bartlett Factor score. Values in bracket are for Annak Broilers.

IV. DISCUSSION

The High VIF values (12.3 and 14.0) obtained in the first set of regression equations confirmed multicollinearity problem among the linear measurements within the two data sets and could lead to inaccurate interpretations of the effect of predictor traits, giving rise to deficiencies [24] in regression models. Thus SD and variances of estimated coefficients in the equations were inflated, giving rise to values far from the true range, and resulting to incorrect conclusions on the relationships between dependent and predictor variables [25]; This relationship would also affect the least squares estimates. The FSR equations reduced multicollinearity by reducing VIF (1.00 and 1.00) statistics, thus removing these problems. DW statistics detect the presence of correlation among prediction errors in regression models and indicate when residuals from a regression are independent. Values ranging from 0 to 2, indicate no autocorrelation among the residuals at the upper threshold. The above procedure maintained stable DW in AA strain from 1.94 to 1.90 and recorded a decrease in AN strain from 1.97 and 1.48. This was interpreted to mean that traits involved have no serial correlation among residual since the values of DW were all close to 2. The R² obtained in our models were considered stable when compared with values obtained in the normal regression. This was an indication of stability of the values of residuals. The R², though lower, showed that the final BFS equations could estimate above 92 % of variability in the data set from respective strain. It also indicated the stability existing among predictors in the models. The t-statistics examine difference between two population means, and is used when variances of two normal distributions are unknown and when an experiment uses small sample size [http://www.investopedia.com/terms/t/t-test.asp]. The higher t-

statistics values obtained indicated improvement and significance (P<0.05) of the regression coefficients. The BFSR procedure employed in this study accounted for 92-94% of variations in data. Total variability extracted in both strains was higher than previous studies on use of FSR. Reference [26] used FS to study the production variables of layer chickens on the deep-litter and obtained 3 orthogonal factors that predicted 70.58% of variability in data. Reference [8] extracted two PC in the Nigerian chicken explaining 66.4% of the total variation in the original variables. By employing Factors Scores extracted from composite traits (trait combinations), we build body conformation index for selection which would yield optimum values for each strain. Thus ADL, LGL, TCF and BRW in AA; and BRW, LGL and HT in AN were measurements that best describe the unique body conformation of both strains in the two factors. Using above BFS regression models we could select for Body-length (ADL, 1.25, BF1) and Meatiness (LGL, 1.21, BF2) in AA strain; and for Breast-size (BRW, 0.7, FS1) and height (HT, 1.81, FS2) in AN strain. This result also revealed that the body conformation traits and indices of broilers could differ as extracted Factors for each strain were primarily loaded by different combinations of body traits. This could be due to different combinations of genes exerting differing influences on different body traits either by correlation and or pleiotropy. The use of BFSR secured improvement over the normal regression since only the common factors had impact on FS. The sum of squared components for the "error" factors or the unique factors across the set of variables was minimized, and the component FS were highly correlated to their corresponding factors and not with other factors. The advantage of BSF over other methods of factor extraction is that it produced unbiased estimates of the true factor scores [27] as these scores were produced by using maximum likelihood

estimates [5]. A single index obtained from regressed model for a bird in respective strain, would be similar to the best linear unbiased prediction (BLUP) of an individual's breeding value on all the sources of information [28], and therefore could be closely related to the population composite breeding value [29]. Thus for an improvement program, individuals with index above a threshold value could be selected while those below could be graded for sales. The advantage of this approach to selection would not be in the reduction of measured traits but in the factoring of traits into groups of inter-correlated traits that are orthogonal to other traits in other factors. This approach overcomes the unrealistic assumptions of equal economic importance of traits, equal heritabilities and variances of traits, and zero correlation among traits, in the traditional index method [29]. All biological traits are of unequal economic importance, unequal heritabilities and variances, and are correlated in nature. Another advantage of this approach to selection is that the effectiveness of equation is expected to

increase with the number of Factors extracted, as obtained in the traditional index method. Also there would be no need for consideration of selection intensity when using BFSR equation.

V. CONCLUSION

BFSR procedure yielded standardized MLR equations for predicting and comparing composite consumer preference traits, for identifying body conformation in broilers; and that could be used for selecting within and discriminating between strains in similar environments.

ACKNOWLEDGMENT

“O. M. A. Jesuyon thanks Dr A. E. Salako for his moral support during the period of this research work.”

REFERENCES

- [1] Y. A. Geidam, M. M. Bukar, and A. G. Ambali, Chick quality control: a key to sustainable poultry production in Nigeria, *Nigerian Veterinary Journal*. 27 (2), 2006, pp. 1-6.
- [2] B. A. Ajayi, and S. O. Oseni, Morphological characterisation and principal component analysis of body dimensions in Nigerian population of adult rabbits. *Proceedings 10th World Rabbit Congress*. September 3 – 6, Sharm El- Sheikh, Egypt, 2012, pp. 229 – 233.
- [3] J. D. Latshaw and B. L. Bishop, Estimating body weight and body composition of chickens by using non-invasive measurements. *Poultry Science*, vol. 80 (7), 2001, pp. 868-873.
- [4] F. O. Ajayi, O. Ejiofor, and M. O. Ironke, Estimation of body weight from linear body measurements in two commercial meat-type chickens. *Global Journal of Agricultural Sciences*, 2008, vol. 7 (1), 2008, pp. 57-59.
- [5] C. DiStefano, M. Zhu, and D. Mindrila, Understanding and Using Factor Scores: Considerations for the Applied Researcher, *Practical Assessment, Research and Evaluation*, 14 (20): 2009, pp. 1-10. ISSN 1531-7714.
- [6] A. A. Skrondal, and P. Laake, Regression Among Factor Scores. *Psychometrika*, 66 (4), 2001, pp. 563 –576.
- [7] K. O. Adekoya, U. D. Abazuh, K. A. Bankole, and B. O. Oboh, A Comparative Morphological Characterization of Three Local and Three Exotic Chicken Types in Nigeria. *Unilag Journal of Medicine, Science and Technology*, [S.l.], Vol 1 (1) Aug. 2015.
- [8] S. S. A. Egena, A. T. Ijaiya, D. M. Ogah, and V. E. Aya, Principal component analysis of body measurements in a population of indigenous Nigerian chickens raised under extensive management system, *Slovak Journal Animal Science*, 47, (2): 2014, pp. 77-82.
- [9] I. Udeh, Use of factor scores for determining the relationship between body measurements and semen traits of cocks. *Open Journal of Animal Sciences*, 2 (1), 2012, pp. 41-44.
- [10] I. Udeh, and C. C. Ogbu. Principal component analysis of Body measurements in three strains of broiler chicken, *Science World Journal*, Vol 6:2, 2011, pp. 11-14.
- [11] A. Yakubu, D. Kuje, and M. Okpeku, Principal Components as Measures of Size and Shape in Nigerian Indigenous Chickens. *Thai Journal of Agricultural Science*. 42(3): 2009a, pp. 167-176
- [12] D. M. Ogah, Assessing size and conformation of the body of Nigerian indigenous turkey. *Slovak J. Anim. Sci.*, 44 (1), 2011, pp. 21-27.
- [26] O. M. A. Jesuyon, Multivariate analysis of production variables of layer chickens on deep-litter in the humid tropics, *Stocarstvo*, 68:2014 (1), 2014, pp. 19-23. UDK 636. 083.14. Accessed October 2015.
- [27] S. L. Hershberger, Factor scores. In: B. S. Everitt and D. C. Howell (Eds.), *Encyclopedia of Statistics in Behavioral Science*, New York: John Wiley, 2005, pp. 636-644.
- [13] D. M. Ogah, A. A. Alaga, and M. O. Momoh, Use of Factor Analysis Scores in Multiple Regression Model for Estimation of Body Weight from Some Body Measurement in Muscovy Duck, *International Journal of Poultry Science*, 8 (11): 2009, pp. 1107-1111.
- [14] A. Yakubu, K. O. Idahor, and Y. Agade, Using factor scores in Multiple linear regression for predicting the carcass weight of broiler chickens using body measurements, *Revista Cientifica UDO Agricola* 9 (4), 2009b, pp. 963 – 967.
- [15] E. Eydurán, K. Karakus, S. Karakus, and F. Cengiz, Usage of factor scores for determining relationships among body weight and some body measurements. *Bulgarian Journal of Agricultural Science*. 15 (4), 2009, pp. 373-377.
- [16] K. A. Shahin, A. M. Soliman, A. E. Moukhtar, Sources of shared variability for the Egyptian buffalo body shape (conformation), *Livest. Prod. Sci.* 36, 1993, pp. 323–334.
- [17] K. A. Shahin, Analysis of muscle and bone weight variation in an Egyptian strain of Pekin ducklings, *Ann. Zootech.* 45, 1996a, pp. 173–184.
- [18] K. A. Shahin, Sources of shared variability in muscle and bone weight distribution and estimation of carcass meatiness and bone utilizing orthogonal carcass traits derived from factor analysis in Japanese quail, *Ann. Zootech.* 46, 1997, pp. 175–183.
- [19] K. A. Shahin, Sources of shared variability in muscle and fat weight distribution in Pekin ducklings, *Ann. Zootech.* 48: 1999a, pp. 59–66.
- [20] K. A. Shahin, Sources of shared variability in meat weight distribution and conformation in Pekin ducklings, *Ann. Zootech.* 48, 1999b, pp. 143–150.
- [21] K. A. Shahin, Selection indexes using live measurements or their varimax rotated factors for improving meat weight distribution – Application on carcasses of Pekin ducks, *Arch. Geflügelkd.* 60: 1996b, pp. 103–108.
- [22] SPSS. Statistical Package for Social Sciences (SPSS) for windows. SPSS Inc., 17.0. 2007. 444 Michigan Avenue, Chicago, 2004, IL60611.
- [23] S. Sharma, *Applied Multivariate Techniques*. Wiley. ISBN: 978-0-471-31064-8, 1996, pp. 512.
- [24] Mendes, M. 2009. Multiple linear regression models based on principal component scores to predict slaughter weight of broiler, *Arch. Geflügelk.*, 73 (2), pp. 139–144.
- [25] S. I. V. Sousa, F. G. Martins, M. C. M. Alvim-Ferraz, M. C. Pereira, Multiple Linear Regression and artificial neural Networks based on Principal Components to predict ozone concentrations, *Environmental Modelling and Software*, 22, 2007 pp. 97-103. In Mendes, 2009.
- [28] D. S. Falconer, and T. F. C. Mackay, *Introduction to quantitative genetics*. 4th edition. Longman group Ltd. Essex, England. ISBN: 0-321-04347-2. 1996, pp. 240-241.
- [29] F. H. Khan, and A. Singh, *Principles of genetics and breeding*. Jaypee Brothers Medical Publishers (P) ltd. New Delhi, India. 2002, pp. 209-213. Ch. 22.