



Prediction of body weight of mixed breeds of pigs in Nigeria through morpho-biometric traits using classification and regression tree models

I. Mallam^{1*}, A. Yakubu², N.P. Achi³

¹Department of Animal Science, Kaduna State University, Kafanchan Campus, Kaduna State, Nigeria

²Department of Animal Science, Faculty of Agriculture, Nasarawa State University, Keffi, Shabu-Lafia, Campus, P.M.B. 135, Lafia, Nasarawa State, Nigeria

³Dairy Research Programme, National Animal Production Research Institute/Ahmadu Bello University, Shika, Zaria

(Manuscript received 18 February 2024; accepted for publication 2 September 2024)

Abstract. *The study was conducted to predict the body weight of mixed breeds of pigs in Nigeria through morpho-biometric traits (body length, chest girth, height at withers, ear length, head length, foreleg length, and hind leg length) using classification and regression tree models. The data were produced using 500 randomly selected mixed breeds of pigs from various farms in five Local Government Areas of Kaduna State, North West Nigeria. The collected data were analysed using the Statistical Package for Social Sciences (SPSS, 2016). Body weight correlated well with morphometric characteristics except with foreleg length, which had a low correlation and no significant ($P>0.05$) difference. Two body dimensions were shown to be more effective in predicting the body weight of the mixed-breeds based on the significance of the independent variables: chest girth and body length. The largest dividing variable was determined to be chest girth, which explained roughly 88.60 % of the difference in body weight. The decision tree model revealed that pigs with chest girth or chest circumference greater than 76.00 cm are expected to have a higher body weight, which livestock producers and researchers could use to determine the feed amount, drug dose, and market price of an animal, as well as the management, selection, and genetic improvement of mixed breeds of pigs in Nigeria.*

Keywords: Prediction, body weight, pigs, morpho-biometric, regression tree, model

Introduction

The pig industry plays a key role in global agriculture and contributes significantly to the supply of meat products for human consumption. As the demand for pork continues to grow, effective and accurate methods of predicting pig body weight are becoming essential for efficient management and optimization of production. Mixed breeds of pigs, which are the result of

different genetic backgrounds, present a unique challenge in predicting body weight due to the inherent variability in their morphometric traits. In recent years, the integration of advanced statistical and machine learning techniques has provided valuable tools for predicting and understanding complex biological phenomena. Among these methods, classification and regression trees (CART) have come to the fore for their ability to discern complex relationships within datasets. This

*e-mail: mallamiliya2011@gmail.com

study seeks to investigate the potential of CART models in predicting the body weight of mixed-breed pigs based on a comprehensive analysis of morphometric traits (Stygar et al., 2018).

Morphometric traits were found to be important predictors of body weight in pigs. Morphometric traits, including measurements of body dimensions and proportions, serve as basic indicators of individual pig growth and development (Kadirvel et al., 2023). Classification and Regression Tree models have been used to predict body weight in pigs, and by leveraging the power of CART models, which work by recursively partitioning data into subsets based on predictor variables, we aim to uncover complex patterns and relationships that may otherwise be difficult to discern with traditional statistical approaches (Kadirvel et al., 2023).

The significance of this study goes beyond the immediate scope of pig farming. Accurate predictions of body weight can increase overall production efficiency and help farmers make informed decisions regarding feed management, husbandry strategies, and resource allocation. In addition, research can intersect genetics, morphometrics, and growth patterns in mixed-breed populations, supporting advances in animal science and breeding practices.

Therefore, the objective of the study was to apply regression tree analysis to predict the body weight (BW) of mixed-breed of pigs through morpho-biometric traits using classification and regression tree models.

Material and methods

Study area and experimental animals

The data used for the study were collected in five local government areas within Kaduna State, namely Kaura, Chikun, Jaba, Zango-Kataf, and Jema'a local government areas. A total of 500 pigs, representing different breeds including Yorkshire, Duroc, Large White, and Hampshire, were measured for various parameters, with a sample size of 100 pigs selected from each of the mentioned local government areas. The pigs were reared semi-intensive and extensively.

Data collection and measurements

The parameters measured were body weight, body length, chest girth, and height at wither; ear length, head length, foreleg length, and hind leg length were measured. Body Weight estimation was carried out using a hanging scale to determine the weight of each animal. Body length was measured as the distance between the occipital protuberance and the tail drop, ensuring precise alignment during measurement. Chest girth was determined as the circumference of the chest just behind the forelimbs, meticulously recorded using a measuring tape. Height at wither was measured as the distance between the most dorsal point of the withers and the ground, ensuring a consistent angle of measurement to maintain accuracy. Amongst other parameters measured were: Ear length, head length, foreleg length, and hind leg length, which were measured using specific tools: ear length and head length were assessed using a measuring tape, while fore leg length and hind leg length were measured using a calibrated wooden calliper. To minimize variations in measurements, all assessments were consistently carried out by the same person, ensuring accuracy and reducing potential discrepancies arising from inter-personnel variations.

Measurement of body weight and biometric characteristics of mixed breeds goats were calculated using Pearson's moment correlation coefficients. Body weight and body measurements were compared to the primary distribution to determine the importance of the variables. The emphasis was on the improvement measure assigned to each variable as a surrogate for the primary distribution (Yohannes and Hoddinott, 2006). Values of improvements were aggregated at each node and scaled to the nearest best performing variable. The variable with the highest sum of improvements had a score of 100, and all other variables had scores from 0 to 100 (Banerjee et al., 2008). The tree building procedure starts by dividing a sample into binary nodes based on a very simple question: $X \leq d$? Where X is a variable in the data set, and d is a real number (Bevilacqua et al., 2003). The root node is the starting point for all observations. The goal is to identify a rule that first divides the observations and creates a group

(or binary node) that is internally homogeneous (Yohannes et al., 2006).

The tree building process continued until it became impossible. The maximum tree value was obtained after the tree reached a maximum dimension. 10-fold cross-validation was used as an error estimation method; this was to provide estimates of the future prediction error for each sub-tree (Camdeviren et al., 2005). The statistical package employed in the analysis was Statistical Package for Social Sciences (SPSS, 2016).

Classification and regression tree (CART) algorithm

Classification and Regression Tree algorithm is one of the tree decision techniques. It was first introduced in 1984 by Breiman et al. (1984). Classification and Regression Tree is widely used in animal industry because of its simplicity, ease of use, and visualization. Classification and Regression Tree was performed according to the following methods: Eydurán et al. (2017); Tyasi et al. (2021) to estimate the dependent variable body weight from 7 biometric traits: body length, chest girth, height at wither, ear length, head length, foreleg length and hind leg length. For each node of CART, the estimated error of prediction explained the dependent variable variation as follows:

$$S^2_x = (1 - S^2_e) \times 100$$

$$S^2_e = \text{risk value}/S^2_y$$

Where in detail:

$$S^2_x = \text{explained variation,}$$

$$S^2_e = \text{unexplained variation}$$

$$S^2_y = \text{variance of the root node (standard deviation of the root node)}^2.$$

Data analysis

The statistical analysis was performed using IBM SPSS (2016) software. The probability of significance was set at 5% and the probability of highly significant was set at 1% between traits. The descriptive statistics used were the mean and standard error, the standard deviation, the minimum

and maximum, and the coefficient of variation. The Pearson correlation was used to predict the relationship between all measured traits. The CART was calculated to create a model.

Results and discussion

Descriptive statistics of body weight (g) and body dimensions (cm) of mixed breeds of growing pigs is presented in Table 1. Each trait was presented with its mean, standard errors, standard deviations, coefficients of variation. In the case of body weight, the average weight stands at approximately 37800 ± 1457.681 g, with the least coefficient of variation = 8.63%. This suggests a moderate level of variability among the observed pigs regarding their weight. Moving to body dimensions such as body length, chest girth, height at wither, ear length, head length, fore leg length, and hind leg length, they exhibit varying levels of standard deviations and coefficients of variation. Average body length is 76.37 ± 1.02 cm and had the coefficient of variation = 29.96%. This indicates a broader range of lengths within the pigs' population. The ear length had the highest coefficient of variation = 33.88% and suggesting more variability among the pigs for the traits compared to other morpho-biometric traits. The data highlights diverse ranges and variability in these measured dimensions among the mixed breeds of growing pigs. These variations may be influenced by genetic diversity, environmental factors, or a combination thereof, emphasizing potential opportunities for genetic manipulation and improvement to optimize desired traits in pig breeding programmes.

The body weight (37800.00 ± 1457.68 g) obtained in this current study was higher than the value of body weight (13140 ± 230 g) reported by Akporhwarho and Iriakpe (2021). The differences could be due to breed, age and some environmental factors, although the values for body length and ear length in this study are within the range reported by the authors.

These recent mean values followed the same trends as those reported by Fayeye et al. (2006); Ogah et al. (2009). The coefficient of variation for body weight and body trait measurements ranged

Table 1. Descriptive statistics of body weight (g) and body dimensions (cm) of mixed breeds of growing pigs

Parameters	Mean± SE	SD	CV (%)
BWT(g)	37800.00 ± 1457.68	3262.73	8.63
BL(cm)	76.37 ± 1.02	22.89	29.96
CG(cm)	81.67 ± 9.49	21.24	26.01
HW(cm)	58.18 ± 0.75	16.85	28.97
EL(cm)	20.03 ± 0.30	6.79	33.88
HDL(cm)	27.91 ± 0.30	6.80	24.36
FL(cm)	39.32 ± 4.34	9.72	24.72
HL (cm)	41.76 ± 0.47	10.54	25.25

SE: Standard error, CV: Coefficient of variation, SD: Standard deviation, BWT: Body weight, BL: Body length CG: Chest girth HW: Height at withers, EL: Ear length, HDL: Head length, FL: Foreleg length, HL: Hind leg length

from 8.63–33.88 and this indicates that body characteristics and body weights are good reliable indices of body size (Akporthuaro and Iriakpe, 2021).

Table 2 presents the phenotypic correlations between body weight and various biometric traits within a mixed breed of growing pigs. Starting with body weight, it exhibits strong positive correlations with certain biometric traits. The result shows a high correlation with withers height at $r=0.813^{**}$ and a notably strong correlation with hind leg length at $r=0.839^{**}$. These correlations are statistically significant, denoted by ** indicating a high level of significance at $p<0.01$. Body length also demonstrates considerable positive correlations with various traits. It exhibits a strong correlation with withers height at $r=0.899^{**}$ and hind leg length at $r=0.896^{**}$, both highly significant. Chest girth, however, displays weaker correlations with other traits, mostly showing non-significant relationships. It demonstrates negligible correlations with the other measured traits, suggesting a lack of substantial association with body weight or other dimensions in this context. The correlations among other measured traits like ear length, head length, foreleg length, and hind leg length vary in strength. These traits exhibit moderate to strong positive correlations with certain other traits, indicating associations between

these morphometric measurements. The observed correlations between body weight and various biometric traits suggest relationships that can inform insights into the pig's physical development. Traits such as withers height and hind leg length appear particularly influential in relation to body weight, potentially serving as reliable indicators or predictors of overall body weight in these growing pigs. The positive correlation between body weight and linear body measurements are similar to the report of Akanno and Ibe (2006) who observed the highest correlation between body weight and shoulder-to-tail length. Similarly, Yakubu (2009) observed that body weight to withers height correlation was (0.95; $p<0.01$) in West African Dwarf goat. Mallam et al. (2018) obtained highest correlation between body weight and head-to-shoulder in New Zealand white rabbits and highest correlation between body weight and body length in chinchilla rabbits.

The coefficients of correlations in this study are comparable to those reported earlier by Yakubu and Ayoade (2009); Yakubu (2010). The high correlation coefficients between body weight and some linear measurements such as body length, height at withers, ear length, head length and hind leg length in this study show that an increase in one body weight will result in a corresponding increase in body size. Consequently, the relationship

Table 2. Phenotypic correlation of body weight and biometric traits of mixed breeds of growing pigs

	BWT(g)	BL(cm)	CG(cm)	HW(cm)	EL(cm)	HDL(cm)	FL(cm)	HL (cm)
BWT(g)	1	0.935**	0.985**	0.813**	0.589**	0.724**	0.050 ^{NS}	0.839**
BL(cm)		1	0.052 ^{NS}	0.899**	0.669**	0.814**	0.031 ^{NS}	0.896**
CG(cm)			1	0.049 ^{NS}	0.007 ^{NS}	-0.030 ^{NS}	-0.001 ^{NS}	0.054 ^{NS}
HW(cm)				1	0.682**	0.841**	0.040 ^{NS}	0.829**
EL(cm)					1	0.786**	0.025 ^{NS}	0.742**
HL(cm)						1	0.042 ^{NS}	0.847**
FL(cm)							1	0.047 ^{NS}
HL (cm)								1

BWT: Body weight; BL: Body length; CG: Chest girth; HW: Height at withers; EL: Ear length; HDL: Head length; FL: Foreleg length; HL: Hind leg length, *=P<0.05, **=P<0.01

existing between body weight and morphometric characters implied that the combination of these morphometric characters can be used to estimate live weight of pigs in places where weighing scales are not available. This, therefore, produces a basis for the genetic manipulation and improvement of the two breeds of pigs in Nigeria.

The importance of each predictor variable in

predicting the body weight of mixed breeds of growing pigs is shown in Table 3. Chest girth was found as the most important variable (100%), followed by body length (97.9%), hind leg length (73.9%), fore leg length (69.4%), height at withers (65.8%), head length (45.4%) and ear length (9.1%). The risk value (estimated) obtained was 121800000 and the standard error found was 25050000.

Table 3. The importance of each predictor variable in predicting the body weight of mixed breeds of growing pigs

Independent Variable Importance			Risk value (estimated)	Std. Error
Independent Variable	Importance	Normalized Importance	121800000	25050000
CG(cm)	942700000	100.0%		
BL(cm)	922700000	97.9%		
HL (cm)	695400000	73.9%		
FL(cm)	653900000	69.4%		
HW(cm)	620600000	65.8%		
HL(cm)	427700000	45.4%		
EL(cm)	85500000	9.1%		

BL: Body length; CG: Chest girth; HW: Height at withers; EL: Ear length; HDL: Head length; FL: Foreleg length; HL: Hind leg length, Std: Standard deviation

The chest girth obtained in this current research as the most important independent variable for predicting body weight of pig is similar with the report of Mallam et al. (2023) who reported chest

girth as the most important independent variable for predicting body weight of Nigerian non-descript goats from morphometric traits using classification and regression tree model. This simply means that

mixed breeds of pigs with higher chest girth will have more body weight than those with lower chest girths. Similarly, Hizli et al. (2022) reported that flock type had 100.00% normalized importance, followed by age of dam (70.80%), birth type (52.21%) and gender (51.70%), respectively, as factors affecting weaning weight of Awassi Lamb.

Figure 1 shows the classification and regression tree model (CART). The model reflects the most morpho-biometric traits that can affect the body weight of mixed breeds of pigs in Nigeria. Node 0 (37747.900 ± 32659.444 g; n = 500) was divided according to chest girth into two subgroups as node 1 ≤ 100 cm (28089.863±18098.222 g; n = 439), node 2 less than 100 cm (107254.098 ± 29760.520 g; n = 61). Node 1 was further subdivided into node 3 and 4. Node 3 ≤ 76 cm (16378.421±7816.169 g; n=285) and node 4 had >76 cm (49763.636±9796.243 g; n=154). Node 3 further subdivided into node 5 and 6 while node 4 further subdivided into 7 and 8. Node 5 had ≤ 60 cm (11787.655±4170.969 g; n=192). Node 6, >60 cm (25856.129±4205.477 g; n=93). Node 7, ≤ 88 cm (42283.951±6047.178 g; n=81). Node 8 had >88 cm (58063.014±5524.655 g; n=73). Node 9 was further subdivided into 9 and 10. Node 9 had ≤ 52 cm (9184.797±2253.250 g; n=123) and node 10 had >52 cm (16427.536±2378.326 g; n=69).

Here, six terminal nodes (nodes 2, 6, 7, 8, 9 and 10) were formed. Node 0, also known as root node, holds descriptive data about the weight of the pigs. The mean body weight in node 2 was higher than the other nodes; however, this data were not sufficient to make a determination. As nodes (2), (6), (7), (8), (9) and (10) were not sub-grouped, they can be considered as homogeneous (terminal) nodes. Node 2 provides a better prediction weight (107254.098±29760.520 g) with a second variance higher (29760.520)² = 885688555.67

In the current study, the root node variance or dependent variable was:

$$S_y^2 = (32659.444)^2 = 1066639151.7513$$

The cause of the unexplained fluctuation in body weight was determined to be:

$$S_e^2 = \text{risk value (estimated)} \div S_y^2 = 121800000 \div 1066639151.7513 = 0.114$$

The regression model's explanation for the

change in the dependent parameter (body weight) was:

$S_y^2 = 1 - S_e^2 = 1 - 0.114 = 0.886 = 88.60\%$. In this study, an algorithm was used to create a model that predicted how much body weight would be gained or lost over time.

The classification and regression tree model in the present study for algorithm was used to create a model that predicts how much body weight would be gained or lost from chest girth and body length of mixed breed of pigs. The current findings indicated that chest girth and body length explained 88.60 % of the variation in body weight of mixed breed of pigs. Based on the model developed in this study, chest girth plays the most significant role in body weight, followed by body length. The risk value indicates the variation within the nodes. This can also be used as the model fitness criterion. The risk value shows the difference within the nodes. The model fitness criterion can also be used. The optimal tree in this study explains about 88.60% of the variation in body weight of the mixed breed of pigs. This means that body weight estimates based on measurements are accurate (Yakubu, 2012, Mallam et al., 2023). The resultant model is easy to understand by people. Afolayan et al. (2006) reported on the use of regression tree methodology for the prediction of slaughter weight of the broiler chickens. Similarly, report by Fourie et al. (2002) stated that chest girth was the primary determinant of body weight for Dorper ram; it also played a significant role in the selection index.

Live body weight is a measurement used in animal husbandry that is important for the successful breeding, nutrition and health management of animals (Mallam et al., 2023). A related study on broiler chickens by Mendes and Akkartal (2009) stated that regression tree methodology was successfully employed to forecast the slaughter weight for broilers. There are many benefits of using a regression tree analysis as opposed to other techniques like multiple regressions or logistic regression and other methods like discriminant and cluster analysis. The results of this study differ slightly from Tyasi et al. (2021), who found that body length (BL) played the most significant effect on body weight, followed by chest girth. Species

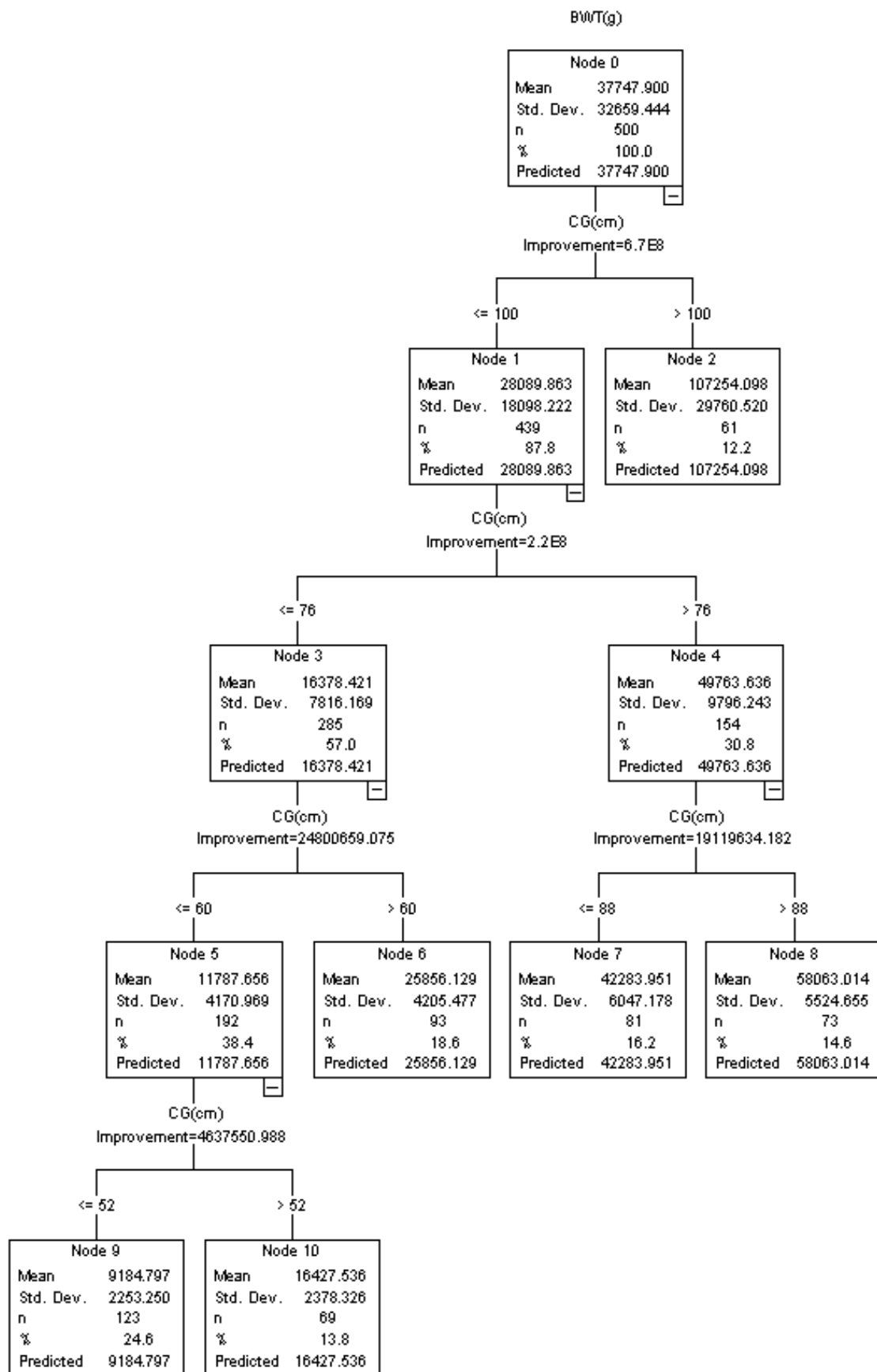


Figure 1. The Classification and Regression Tree Model

differences and other factors may also contribute to body weight variation.

Conclusion and Recommendations

It can be concluded that there is a positive correlation between body weight and all the morpho-biometrical traits measured but low correlation ($r=0.05$) with foreleg length of the mixed breed of pigs in Nigeria. The results of this study suggest that the morpho-biometrical characteristics of the mixed breed of pigs in Nigeria particularly chest girth and body length may be employed as selection criteria to increase their body weight.

As the CART model has a high coefficient of determination, chest girth alone can be used to predict the body weight of pigs in Nigeria. The findings of this study may be useful for livestock producers and researchers to determine the feed quantity and market price of the animal to increase the profitability of the animal farm, selection, and management as weight is the primary factor for animal production.

References

- Afolayan RA, Adeyinka IA and Lakpini CAM**, 2006. The estimation of live weight from body measurements in Yankasa sheep. *Czech Journal of Animal Science*, 51, 343-348.
- Akanno EC and Ibe SN**, 2006. Prediction of body weight of the domestic rabbit at different stages of growth using linear body measurements. *Nigerian Journal of Animal Production*, 33, 3-8.
- Akporhualho PO and Iriakpe O**, 2021. Assessment of body weight and morphological traits of two breed of grower pigs using principal component analysis. *Nigerian Journal of Animal Production*, 48, 1 - 11. <https://doi.org/10.51791/njap.v48i5.3182>.
- Banerjee AK, Arora N and Murty USN**, 2008. Classification and regression tree (CART) analysis for deriving variable importance of parameters influencing average flexibility of CaMK Kinase Family. *Electronic Journal of Biology*, 4, 27-33.
- Bevilacqua M, Braglia M and Montanari R**, 2003. The classification and regression tree approach to pump rate analysis. *Reliability Engineering and System Safety*, 79, 59-67.
- Breiman L, Friedman J, Stone CJ and Olshen RA**, 1984. Classification and regression trees. CRC Press, USA.
- Camdeviren H, Mehmet M, Ozkan MM, Toros F, Sasmaz T and Oner S**, 2005. Determination of depression risk factors in children and adolescents by regression tree methodology. *Acta Medica Okayama*, 59, 19-26.
- Eyduran E, Zaborski D, Waheed A, Celik S, Karadas K and Grzesiak W**, 2017. Comparison of the Predictive Capabilities of Several Data Mining Algorithms and Multiple Linear Regression in the Prediction of Body Weight using Body Measurements in the Indigenous Beetal Goat of Pakistan. *Pakistan Journal of Zoology*, 49, 257-265.
- Fayeye TR, Ayorinde KL, Ojo V and Adesina OM**, 2006. Frequency and influence of some major genes on body weight and body size parameters of Nigerian local chickens. *Livestock Research for Rural Development*, 18, Article #37. Retrieved February 17, 2024. <http://www.lrrd.org/lrrd18/3/faye18037.htm>.
- Fourie PJ, Naser FWC, Olivier JJ and van der Westhuizen**, 2002. Relationship between production performance, visual appraisal and body measurements of young Dorper rams. *South African Journal of Animal Science*, 32, 256-262.
- Hizli1 H, Takma C and Ergül S**, 2022. Application of Classification and Regression Tree (CRT) Method for Predicting the Some Environmental Factors Affecting Weaning Weight of Awassi Lamb. *Journal of Agricultural Faculty of Gaziosmanpasa University*, 39, 185-190. doi:10.55507/gopzfd.1169234.
- Kadirvel G, Devi YS, Naskar S, Bujarbaruah KM, Khargariah G, Banik S, Singh NS and Gonmei C**, 2023. Performance of crossbred pigs with indigenous and Hampshire inheritance under a smallholder production system in the Eastern Himalayan hill region. *Frontier Genetics*, 14, 102554. doi: 10.3389/fgene.2023.1042554.
- Mallam I, Hussaini YI, Alhassan ID, Ngedu EO, Kehinde WH and Gugong V**, 2023. Prediction of Body Weight of Nigerian Non-Descript Goats from Morphometric Traits Using Classification

and Regression Tree Model. FUDMA Journal of Agriculture and Agricultural Technology, 9, 109-116.

Mallam I, Kabir M, Nwagu BI, Achi NP, Achi JN, Alao RO and John PA, 2018. Influence of Genotype on Post-Weaning Growth Performance of Domestic Rabbits. Nigerian Journal of Animal Science (NJAS), 20, 17-25.

Mendes M and Akkartal E, 2009. Regression tree analysis for predicting slaughter weight in broilers. Italian Journal of Animal Science, 8, 615-624.

Ogah MD, Hassan ID and Musa IS, 2009. Path analysis of the relationship between various body measurements and live weight in immature West African Dwarf goats. Analele IBNA, 25, 72-77.

Statistical Package for Social Sciences, 2016. IBM Corporation released 2016. Ibm SPSS, Statistics for windows, version 16.0 armonk. New York: IBM, Corporation SPSS.

Stygar AH, Dolecheck KA and Kristensen AR, 2018. Analyses of body weight patterns in growing pigs: a new view on body weight in pigs for frequent monitoring. Animal, 12, 295-302. <https://doi.org/10.1017/S1751731117001690>.

Tyasi TL, Eyduran E and Celik S, 2021. Comparison of tree-based regression tree methods for predicting live body weight from morphological traits in Hy-line silver brown commercial layer and indigenous Potchefstroom Koekoek breeds raised in South Africa. Tropical Animal Health and Production, 53, 7.

Yakubu A, 2009. Fixing collinearity instability in the estimation of body weight from morpho-biometrical traits of West African dwarf goats. Trakia Journal of Sciences, 7, 61-66.

Yakubu A, 2010. Path coefficient and path analysis of body weight and biometric traits in Yankasa lambs. Slovak Journal of Animal Science, 43, 17-25.

Yakubu A and Ayoade JA, 2009. Application of principal component factor analysis in quantifying size and morphological indices of domestic rabbits. International Journal of Morphology, 27, 1013-2017.

Yohannes Y and Hoddinott J, 2006. Classification and regression trees: An introduction. A technical guide. International Food Research Institute, 2033 K Street, N.W., Washington, D.C., 27.