UDC 575. 630
https://doi.org/10.2298/GENSR2203429Y
Original scientific article

USE OF LINEAR MODELING, MULTIVARIATE ADAPTIVE REGRESSION SPLINES AND DECISION TREES IN BODY WEIGHT PREDICTION IN GOATS

Abdulmojeed YAKUBU[1]*, Ecevit EYDURAN[2], Senol CELIK[3], Juliana O. ISHAYA[1]

[1]Department of Animal Science, Faculty of Agriculture, Nasarawa State University, Keffi, Shabu-Lafia Campus, P.M.B. 135, Lafia, 950101, Nigeria.
[2]Departments of Animal Science and Business Administration, Igdir University, Igdir, Turkey
[3]Department of Animal Science, Faculty of Agriculture, Bingol University, Bingol, Turkey

Yakubu A., E. Eyduran, S. Celik, J. O. Ishaya (2022). *Use of linear modeling, multivariate adaptive regression splines and decision trees in body weight prediction in goats*. - Genetika, Vol 54, No.3, 1429-1445.
Use of robust regression algorithms for better prediction of body weight (BW) is receiving increased attention. The present study therefore aimed at predicting BW from chest circumference, breed and sex of a total of 1,012 goats. The animals comprised 332 matured West African Dwarf (WAD) (197 bucks and 135 does), 374 Red Sokoto (RS) (216 bucks and 158 does) and 306 Sahel (SH) (172 bucks and 134 does) randomly selected in Nasarawa State, north central Nigeria. BW prediction was made using automatic linear modeling (ALM), multivariate adaptive regression splines (MARS), classification and regression tree (CART), chi-square automatic interaction detection (CHAID) and exhaustive CHAID. The predictive ability of each statistical approach was measured using goodness of fit criteria i.e. Pearson's correlation coefficient (r), Coefficient of determination (R^2), Adjusted coefficient of determination (Adj. R^2), Root-mean-square error (RMSE), Mean absolute percentage error (MAPE), Mean absolute deviation (MAD), Global relative approximation error (RAE), Standard deviation ratio (SD_{ratio}), Akaike's information criterion (AIC) and Akaike's information criterion corrected (AICc). Male RS and SH goats had significantly ($P<0.05$) higher BW and CC compared to their female counterparts while in WAD, male goats had significantly ($P<0.05$) higher CC (57.88±0.51 vs. 55.45±0.55). CC was determined to be the trait of paramount importance in BW prediction, as expected. Among the five models, MARS algorithm gave the best fit in BW prediction with r, R^2, Adj. R^2, SD_{ratio}, RMSE, RAE, MAPE, MAD, AIC and AICc values of 0.966, 0.933, 0.932, 0.26, 1.078, 0.045, 3.245, 0.743, 186.0 and 187.0, respectively. The present information may guide the choice of model which may be exploited in the selection and genetic improvement of animals

Corresponding author: Abdulmojeed Yakubu**,** Department of Animal Science, Faculty of Agriculture, Nasarawa State University, Keffi, Shabu-Lafia Campus, P.M.B. 135, Lafia, 950101, Nigeria, email address: abdulmojyak@gmail.com; abdulkubu@nsuk.edu.ng

including feed and health management and marketing purposes, and especially in the identification of the studied breed's standards.

Key words: body weight, goats, modelling, regression algorithms, tropics

INTRODUCTION

Production of goats reared mostly by rural farmers has the potential of improving the livelihoods of poor livestock keepers, reducing poverty and attaining sustainable agriculture and food security (LANDAU, 2017; YUSUF *et al*., 2018). They are easy to raise compared to large animals and can survive in harsh conditions with less investment. Population of goats in Nigeria, south Saharan Africa was estimated at 78,037,077 million (FAOSTAT, 2017). There are basically three goat breeds comprising the Sahel, Red Sokoto and West African Dwarf goats reared in Nigeria. However, the first two are traditionally adapted to the northern axis of the country; the third breed is prevalent in the southern axis where it is known to be trypanotolerant (YAKUBU *et al*., 2016; OSENI *et al*., 2017). Considering the importance of goats under the rural conditions, there is a need for continuous improvement to scale up production in order to increase income and livelihoods of livestock keepers in Nigeria. In this regard, one of the most important traits is body weight (BW). The BW of animals is useful not only to determine growth, but also the carcass, nutritional and health requirements, and correction of tail length (ALVES *et al*., 2019; ECK *et al*., 2019). This trait is highly heritable and its improvement has been reported to be feasible under traditional smallholders' breeding practices (GIZAW *et al*., 2014). However, under village settings, taking BW of animals may be cumbersome due to the non-availability of weighing scales with good precisions. As a result of this, BW is normally predicted from simple linear body measurements (DORANTES-CORONADO *et al*., 2015).

The prediction of BW from morphometric traits is receiving increasing attention (CAMPOS *et al*., 2017; CORA *et al*., 2019; CANUL-SOLIS *et al*., 2020). Chest circumference (CC) is one of the most important morphological traits for the BW prediction (TEMOSO *et al*., 2017; ZERGAW *et al*., 2017; AMEEN and MIKAIL, 2018; HABIB *et al*., 2019). Due to breed and environmental effects, alternative models might be required in the BW prediction of goats (EYDURAN *et al*., 2010; WORKU, 2019). While previous efforts laid emphasis on simple and multiple regression equations in the BW prediction; these have been found to be deficient as a result of poor predictions occasioned by multicollinearity problem (YAKUBU *et al*., 2009; VALSALAN *et al*., 2020). To remove the problem, some earlier researchers adopted to jointly use scores of factor analysis and principal component analysis in multiple linear regression analysis (EYDURAN *et al*., 2013). However, data mining algorithms such as multivariate adaptive regression splines (MARS), automatic linear modeling (ALM), classification and regression tree (CART), chi-square automatic interaction detection (CHAID) and exhaustive CHAID (these last three models permit easy comprehension using graphics) are receiving increased attention in livestock modeling. These models are more robust and are less error-prone in comparison with the traditional ordinary least square method (EYDURAN *et al*., 2017; AKKOL 2018; OLFAZ *et al*., 2018; GORCZYCA *et al*., 2018; YAKUBU *et al*., 2018a; RAD NAROUI *et al*., 2020). These non-parametric algorithms can also efficiently handle large datasets devoid of any ambiguous parametric structure (SONG and LU, 2015). They have been successfully used in predicting BW (CELIK *et al*., 2017; EYDURAN *et al*., 2017), growth responses (AKIN *et al*., 2020), age (CORRON *et al*., 2016), milk yield (EYDURAN et al. 2013), nutritional efficiency and energy expenditure (ZAKERI *et al*., 2010; CANNAS *et al*., 2019) and clinical conditions (RATIVA *et al*., 2018)

In Nigeria, there is a dearth of information on the use of modern analytical techniques in the prediction of BW. Therefore, an attempt was made in the current study to predict BW from CC, breed and sex of goats using MARS, ALM, CART, CHAID and Exhaustive CHAID for better accuracy and utilization.

MATERIALS AND METHODS

Sampling locations

The study was conducted in four randomly selected popular goat markets (Karu, Keffi, Akwanga and Lafia), found in Nasarawa State, north central Nigeria. Nasarawa State is located in the guinea savannah agro-ecological zone of Nigeria and lies between latitudes 7° 52′ N and 8° 56′ N and longitudes 7° 25′ E and 9° 37′ E respectively (LYAM, 2007). The mean annual rainfall is at least 1600 mm and lowest mean monthly relative humidity of not less than 70%. The mean annual maximum temperature varies from 35 to 31 °C all year round while the mean annual minimum is between 23 and 20 °C (YAKUBU *et al*., 2019). The experimental protocol was in accordance with the set guidelines of ARRIVE **(**Animal Research: Reporting of In Vivo Experiments).

Collection of data

Body weight (BW) and CC measurements of 1,012 mature goats of both sexes (8-tooth permanent incisors) were taken. These covered the three Nigerian indigenous goat breeds [332 WAD (197 bucks and 135 does), 374 RS (216 bucks and 158 does) and 306 SH (172 bucks and 134 does]. Before measurements were taken on each animal, the goat keepers were asked about the system of management. Based on this information, only animals that were subjected to the traditional extensive rearing system were sampled. This was to reduce to the bearest minimum errors due to non-uniformity of management system. BW of live animal was taken using a hanging scale with a capacity of 50 kg and an accuracy of 10 g. **CC** was measured just behind the forelimbs using a measuring tape. Both measurements were taken early in the morning before the animals were fed.

Data analysis

In order to test the effect of sex within WAD, RS and SH goats, t-statistic was used. The separation of means has been made using two-sample t test at $P < 0.05$ significance level. The prediction of BW from CC, breed and sex was done using ALM, MARS, CART, CHAID and Exhaustive CHAID algorithms (KOVALCHUK *et al*., 2018; EYDURAN *et al*., 2019a).

MARS is a nonparametric regression technique that approximates a complex nonlinear relationship by a series of spline functions on different intervals of the independent variable. The MARS model (FRIEDMAN, 1991) can be written in form of:

$$y = f(x) + \varepsilon$$

where, ε = error term, $x = (x_1, x_2, \ldots, x_p)^T$ = p number of predictor variables, and y = response variable.

After the removal of each basis function, the MARS model is refitted and each reduced sub-optimal model is tested with the Generalized Cross-Validation (GCV) method to avoid overfitting (AKIN *et al*., 2020) as indicated below:

$$GCV(\lambda) = \frac{SSE}{\left[1 - \frac{M(\lambda)}{n}\right]^2}$$

Where,
n = number of training cases,
M (λ) = penalty function for the complexity of the model containing λ terms.

CART growing method with a graphical representation of output was subjected to cross-validation with 10 sample folds to estimate error as earlier described (ALI *et al*., 2015; YAKUBU, 2012; EYDURAN *et al*., 2017; YAKUBU *et al*., 2018)

CHAID and Exhaustive CHAID were also used to model BW. CHAID is another tree-based model proposed by KASS (1980) with merging, partitioning and stopping stages that recursively uses multi-way splitting procedures to form homogenous subsets using Bonferroni adjustment until the least differences between the predicted and actual values in a response variable are obtained (EYDURAN *et al*., 2016b; CELIK *et al*., 2017). The Exhaustive CHAID, as a modification of CHAID algorithm, applies a more detailed merging and testing of predictor variables (CELIK *et al*., 2017).

The predictive performance of each of ALM, MARS, CART, CHAID and Exhaustive CHAID was assessed using the following model evaluation goodness-of-fit criteria (AHMADI *et al*., 2007; EYDURAN *et al*., 2017; CELIK *et al*., 2017):

Coefficient of determination (R^2):

$$R^2 = 1 - \frac{\sum_{i=1}^{n}(Y_i - \hat{Y}_i)^2}{\sum_{i=1}^{n}(Y_i - \bar{Y})^2}$$

Adjusted coefficient of determination (Adj. R^2):

$$Adj.R^2 = 1 - \frac{\frac{1}{n-k-1}\sum_{i=1}^{n}\left(Y_i - \hat{Y}_i\right)^2}{\frac{1}{n-1}\sum_{i=1}^{n}(Y_i - \bar{Y})^2}$$

Root-mean-square error (*RMSE*):

$$RMSE = \sqrt{\frac{1}{n}\sum_{i=1}^{n}(Y_i - \hat{Y}_i)^2}$$

Mean absolute percentage error (*MAPE*):

$$MAPE = \frac{1}{n}\sum_{i=1}^{n}\left|\frac{Y_i - Y_{ip}}{Y_i}\right| \times 100$$

Mean absolute deviation (*MAD*):

$$MAD = \frac{1}{n}\sum_{i=1}^{n}\left|Y_i - Y_{ip}\right|$$

Global relative approximation error (*RAE*):

$$RAE = \sqrt{\frac{\sum_{i=1}^{n}(y_i - y_{ip})^2}{\sum_{i=1}^{n} y_i^2}}$$

Standard deviation ratio (SD_{ratio}):

$$SD_{ratio} = \frac{s_m}{s_d}$$

Akaike's information criterion (*AIC*):

$$AIC = nlog\left(\frac{RSS}{n}\right) + 2k$$

Akaike's information criterion corrected (*AICc*):

$$AICc = AIC + \frac{2k\,(k+1)}{n-k-1}$$

Where n is the number of cases in a set, *k* is the number of model parameters, Y_i is the observed value of BW, Y_{ip} is the predicted value of BW, *sm* is the standard deviation of model errors, *sd* is the standard deviation of BW, RSS: Residual sum of squares.

MARS algorithm was specified by the earth package (MILBORROW, 2011, 2018) of R Studio (version 3.5.2) program (R CORE TEAM, 2014). The analyses involving the use of ALM, CART, CHAID and Exhaustive CHAID algorithms were carried out using IBM SPSS (2015). To calculate goodness of fit criteria, ehaGoF package was used (EYDURAN, 2019b).

RESULTS

The sex effect on BW of goats in Table 1. reveal that, WAD goats males were not significantly (P>0.05) heavier than females despite the fact that the former had significantly (P<0.05) higher CC (57.88±0.51 vs 55.45±0.55). However, male RS and SH goats had higher BW and CC compared to their female counterparts.

Table 1. Effect of sex on body weight and chest circumference (Means±S.E.) of goats

Breed	Body weight (BW)		Chest circumference (CC)	
	Male	Female	Male	Female
West African Dwarf (WAD)	20.57±0.23[a]	20.14±0.26[a]	57.88±0.51[a]	55.45±0.55[b]
Red Sokoto (RS)	23.69±0.19[a]	22.47±0.19[b]	61.89±0.24[a]	59.90±0.30[b]
Sahel (SH)	28.19±0.23[a]	27.13±0.25[b]	69.92±0.27[a]	68.87±0.32[b]

S.E.= standard error
Means with different superscripts in the same row are significantly different (P<0.05) for each goat breed

Within breed, prediction of BW from CC and sex revealed that CC was the most determinative trait (P<0.01) in WAD, RS and SH goats to predict BW within the scope of ALM algorithm (Table 2). The inclusion of sex especially male in the model was with negligible contribution (0.002) and not significant (P>0.05). However, when the data were pooled for the three goat breeds, the incorporation of CC, WAD and RS in the model were significant (P<0.01) while the inclusion of SH was redundant.

Table 2. Regression coefficients and fractional importance of variables affecting body weight prediction in goats using automatic linear modeling

Model term	Coefficient	Significance (p-value)	Importance
West African Dwarf			
Intercept	-1.389	0.064	
Chest circumference	0.383	0.000	1.000
Red Sokoto			
Intercept	-17.055	0.000	
Chest circumference	0.659	0.000	1.000
Sahel			
Intercept	-25.455	0.000	
Chest circumference	0.764	0.000	0.998
Male	0.259	0.067	0.002
Female	0[a]	-	0.002
Pooled data			
Intercept	-6.646	0.000	
Chest circumference	0.495	0.000	0.982
West African dwarf	-1.107	0.000	0.018
Red Sokoto	-0.387	0.009	0.018
Sahel	0[a]	-	0.018

[a] This coefficient is set to zero because it is redundant

The observed and predicted BW of goats summary statistics are presented in Table 3. The predicted BW mean values using ALM model in WAD (20.399), RS (23.176) and SH (27.727) were the same with their respective observed values. The associated error term "Variance" for predicted and observed in each case was WAD (7.139 vs. 9.876), RS (5.659 vs. 7.158) and SH (7.830 vs. 9.288).

Table 3. Descriptive statistics of the observed and predicted body weight in goats using automatic linear modeling

Body weight	Minimum	Maximum	Mean	Variance
West African Dwarf				
Observed	14.0	25.0	20.399	9.876
Predicted	15.8	25.4	20.399	7.139
Red Sokoto				
Observed	18.00	30.00	23.176	7.158
Predicted	16.29	29.84	23.176	5.659
Sahel				
Observed	22.0	32.8	27.727	9.288
Predicted	20.4	32.1	27.727	7.830

CC positively affected BW of goats when it was >58 cm with a coefficient of 0.603 for only 4th term (Table 4). All the coefficients were found very significant (P<0.001). CC * Breed interaction revealed that when CC exceeded 68 cm in SH goats, it tended to increase BW more than other interactions with a coefficient of 2.090. In WAD goats, when CC was >60 cm and having a coefficient of 1.902, it contributed more to BW compared to CC >48 cm (1.374 coefficient) and CC > 65 cm (1.244 coefficient). RS breed was redundant in MARS model while the influence of sex was negligible.

Table 4. Body weight prediction in goats using multivariate adaptive regression splines

Model	Coefficient	Standard error	t-value	Pr(>|t|)
Intercept	22.664	0.512	44.305	< 2e-16***
BREEDSAHEL	-5.504	0.396	-13.902	< 2e-16***
BREEDWAD	-3.402	0.465	-7.314	5.32e-13***
h(CHESTCIR-58)	0.603	0.058	10.467	< 2e-16***
h(CHESTCIR-67)	-0.342	0.089	-3.860	0.000121***
h(68-CHESTCIR)	-0.199	0.045	-4.387	1.27e-05***
h(68-CHESTCIR)*SEXmale	-0.032	0.008	-4.024	6.15e-05***
h(CHESTCIR-68)*BREEDSAHEL	2.090	0.211	9.899	< 2e-16***
h(70-CHESTCIR)*BREEDSAHEL	0.759	0.075	10.143	< 2e-16***
h(CHESTCIR-70)*BREEDSAHEL	-1.287	0.209	-6.146	1.15e-09***
h(48-CHESTCIR)*BREEDWAD	1.074	0.104	10.333	< 2e-16***
h(CHESTCIR-48)*BREEDWAD	1.374	0.142	9.667	< 2e-16***
h(CHESTCIR-50)*BREEDWAD	-1.244	0.162	-7.655	4.57e-14***
h(CHESTCIR-58)*BREEDWAD	-0.910	0.154	-5.921	4.39e-09***
h(CHESTCIR-60)*BREEDWAD	1.902	0.214	8.870	< 2e-16***
h(CHESTCIR-62)*BREEDWAD	-2.579	0.213	-12.092	< 2e-16***
h(CHESTCIR-65)*BREEDWAD	1.244	0.164	7.586	7.55e-14***

*** Significant at P<0.001

A total of 12 terminal nodes (nodes 3, 10, 13, 14, 15, 16, 17, 18, 19, 20, 21, and 22) were generated in the prediction of BW using CART algorithm (Figure 1). CC was more important than Breed in BW estimation. Among the terminal nodes, the goats with CC >73.500 cm had the heaviest BW of 32.268 kg [variance: $(0.975)^2 = 0.951$] in Node 22 followed by Nodes 21 (30.833 kg), 20 (29.091 kg) and 19 (27.625 kg). However, the lightest weight (16.711 kg) was recorded in goats with CC <= 51.000 cm in node 3. There was Breed and CC interaction where RS goats with CC > 64.000 cm had a weight of 27.307 in node 18, SH and RS goats > 58.500 cm had a weight of 22.213 kg in node 15 while their WAD counterparts had a weight of 20.955 kg in node 16. No significant effect of sex was found in the prediction of BW.

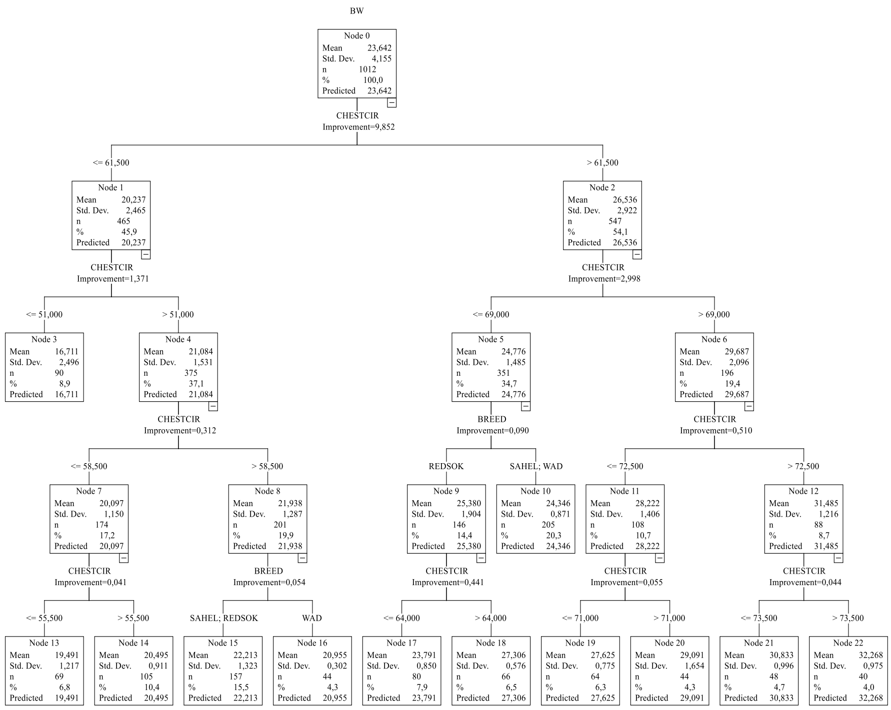

Figure 1. A graphical representation of body weight prediction using CART

In the CHAID analysis, 11 terminal nodes (nodes 2, 3, 5, 6, 7, 8, 9, 10, 11, 12 and 13) were produced (Figure 2). CC was also the trait of paramount importance compared to Breed and Sex in BW estimation. Among the terminal nodes; node 9 with CC >72.000 cm had the best BW of 31.485 kg [variance; $(1.216)^2 = 1.479$) compared to others.

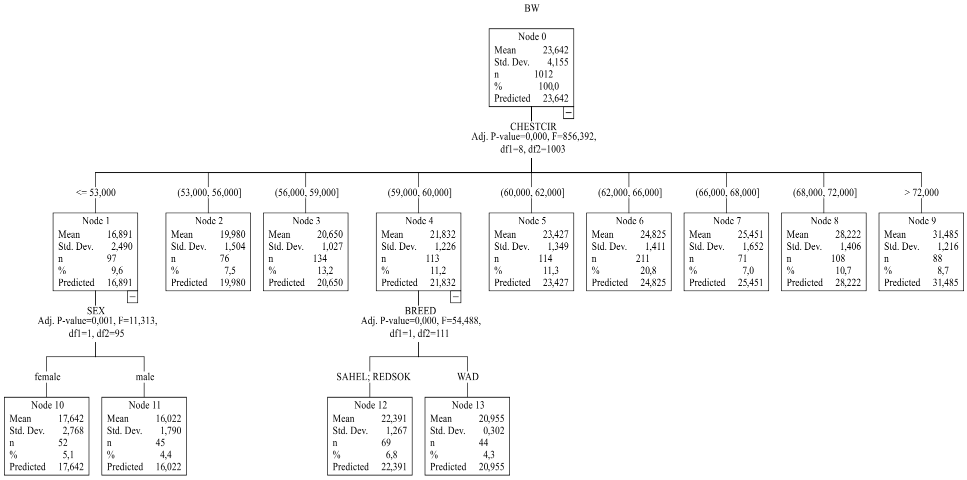

Figure 2. A graphical representation of body weight prediction using CHAID

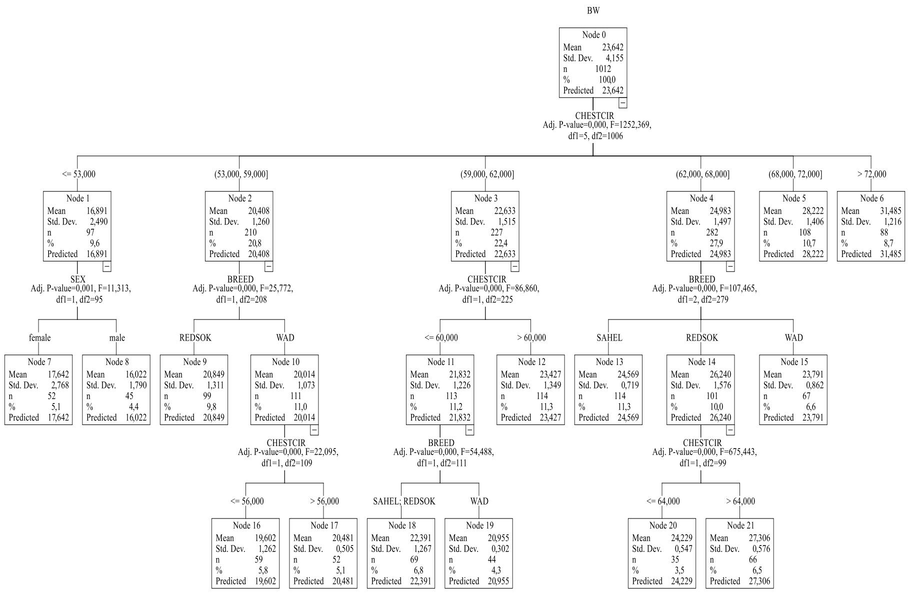

Figure 3. A graphical representation of body weight prediction using Exhaustive CHAID

In the Exhaustive CHAID analysis, 14 terminal nodes (nodes 5, 6, 7, 8, 9, 12, 13, 15, 16, 17, 18, 19, 20 and 21) were generated (Figure 3). CC was also the superior trait compared to Breed and Sex in the estimation of BW. Among the terminal nodes; node 6 with CC >72.000 cm had the best BW of 31.485 kg [variance; $(1.216)^2 = 1.479$) compared to others. The CC, mean BW value and variance are exactly the same with those recorded for CHAID.

Based on the goodness-of-fit criteria employed in the current study, MARS algorithm appeared to be the best in BW predicting BW of goats followed by, CART, Exhaustive CHAID, CHAID and ALM algorithms (Table 5). The r, R^2, Adj. R^2, SD_{ratio}, RMSE, RAE, MAPE, MAD, AIC and AICc values ranged from 0.925-0.966, 0.855-0.933, 0.855-0.932, 0.260-0.381, 1.078-1.584, 0.045-0.060, 0.005-5.065, 0.743-1.219, 186.000-937.194 and 187.000-937.206.

Table 5. Model evaluation criteria for body weight prediction in goats

Algorithms	r	R^2	Adj. R^2	SD_{ratio}	RMSE	RAE	MAPE	MAD	AIC	AICc
ALM	0.925	0.855	0.855	0.381	1.584	0.052	0.005	1.219	937.194	937.206
CHAID	0.938	0.879	0.879	0.348	1.444	0.060	5.065	1.148	749.448	749.472
Exh. CHAID	0.953	0.908	0.908	0.303	1.259	0.052	4.236	0.941	471.810	471.834
CART	0.956	0.915	0.915	0.292	1.214	0.051	3.789	0.833	398.261	398.285
MARS	0.966	0.933	0.932	0.260	1.078	0.045	3.245	0.743	186.0	187.0

DISCUSSION

Worldwide recognition of goats as a veritable supply of meat and milk has been documented (GARCÍA-MUÑIZ *et al*., 2019). The higher average body weight and chest circumference of RS goats compared to their WAD counterparts is consistent with the findings of OKPEKU *et al*., (2011) in goats reared in the southern parts of Nigeria. In a related study, SOWANDE *et al*. (2010) reported average mean values of 20.6 kg (BW) and 66.4 cm (CC) for WAD goats that were 25–36 months old. However, the weight values recorded for WAD goats in the present values are different from the range of 16.4-19.6 kg reported for the same breed in south western part of Nigeria (Olatunji-AKIOYE and ADEYEMO, 2009). Also, under tropical conditions, FAHIM *et al.* (2013) reported for Indian goats greater than 18 months of age mean BW and CC of 18.35±0.24 kg and 61.04±0.38 cm, respectively. Higher BW (32-34.5 kg) and CC (74.7- 77.3 cm) values were recorded for 3-year-old local does in Mexico (DORANTES-CORONADO *et al*., 2015). These differences could be as a result of varying genetic potential, age and environmental conditions. It is expected that animals with genes for higher weight will outperform their counterparts with genes for lower weight. According to RAMOS *et al*. (2019), edaphoclimatic conditions can affect breed performance, hence the need for the rearing of only those breeds that are able to exhibit their genetic potential including body conformation in a particular environment. The superior weight and girth advantage of SH goats to the two other goat breeds in Nigeria could be attributed to the genetic make-up of SH goats as a bigger animal compared to RS and WAD goats. Also, age is an important determinant factor of the weight of an animal. During the growing periods, animals exhibit different body weights at different ages. While lower weights are recorded at birth, higher weights are obtained with age advancement (DAKHLAN *et al*., 2021; TYASI *et al*., 2022). It is therefore imperative to obtain separate prediction models for body weight at each age category. Higher morphometric values were observed in bucks compared to does in the present study. This is consistent with the sexual selection hypothesis (POLAK and FRYNTA, 2010), attributable to delayed sexual maturation and prolonged growth of males. However, ISAAC (2005) reported that male-biased sexual dimorphism in body size is common, but its occurrence certainly is not the exclusive pattern. This was substantiated by the findings of ABD-ALLAH *et al*. (2019) in Egyptian goats.

Several studies have reported chest circumference-based prediction model. VANVANHOSSOU *et al*. (2018) reported that the model incorporating only CC was sufficient enough for BW estimation. Similarly, BEDADA *et al*. (2019) submitted that the contribution to prediction accuracy of BW from CC by other linear body measurements was negligible. The use of CC is important due to its muscle, some fat, and bone structure composition (DORANTES-CORONADO *et al*., 2015). The present information on CC may be exploited by goat producers for feeding and health management, selection and genetic improvement (ABD-ALLAH *et al*., 2019; LATIFI and RAZMKABIR, 2019). This is consistent with the submission of EYDURAN *et al*. (2016a).

The goodness-of-fit criteria are estimated in regression analyses to assess the performance and efficiency of each model. The R^2 and adjusted R^2 of the best model are usually the greatest with the lowest values of RMSE, MAPE MAD and RAE, respectively. For a model to be of a good fit, SD ratio is <0.40; and for it to be of a very good fit, the ratio is <0.10 (ALI *et al*., 2015; EYDURAN *et al*., 2016). In the current study, ALM model had the least predictive ability. Although ALM is said to be more reliable than the traditional linear regression models, its shortcomings include non-inclusion of partitioning/splitting and terms, which would have made it possible for researchers to split data easily into two sets for training and validation (YANG, 2013). MARS and CART appeared to be more reliable in BW prediction. This might not be unconnected with the fact that the underlying idea of MARS modeling is a combinatorial heuristic, which constructs a mathematical model of a system in an evolutionary fashion (KOC and BOZDOGAN, 2015). A unique feature of MARS is that different subregions of its predictors might have different interaction patterns. MARS is also believed to be a generalization of the stepwise linear regression or modification of the CART model for better predictive performance (KOC and BOZDOGAN, 2015).

The set of evaluation criteria used in the current study performed better in comparison with similar set reported in Pakistan goats by CELIK (2019): However, MARS model was the best in the latter with corresponding R^2 (0.91), adjusted R^2 (0.86), RMSE (3.32), SD ratio (0.30), MAPE (8.49), MAD (2.67), RAE (0.09), AIC (402) and AICc (451), followed by CHAID, Exhaustive CHAID and CART. EYDURAN *et al*. (2017) also reported in Beetal goats of Pakistan RMSE (4.4687 vs. 4.1569), MAPE (8.1208 vs. 7.2946) MAD (3.3251 vs. 2.9904), RAE (0.1000 vs. 0.0930), SD ratio (0.5706 vs. 0.5308) and AIC (619.81 vs. 594.16) values for CART and CHAID models, respectively. Using CART algorithm in dogs, CELIK and YILMAZ, (2018) obtained 0.6889 R^2, 0.6810 Adj. R^2, 0.5549 SD ratio and 1.1802 RMSE. However, MARS recorded the best R^2 = 0.9193, Adj. R^2=0.8983, SD ratio = 0.2840 and RMSE = 0.6041. In another study, YAKUBU (2012) and YAKUBU *et al*. (2018b) used CART to successfully predict BW of rams and egg number of laying birds. However, the varying model evaluation values observed may be attributed to breed of animal, sex, age, physiological status, production systems, environment, and sensitivity of the models (YAKUBU, 2010; YAKUBU and MOHAMMED, 2012; MOKOENA *et al*., 2022): As a result of these influencing factors, it is possible that each climatic region or race might have a particular prediction model(s) best suited to it.

CONCLUSION

The study revealed that sexual dimorphism was largely in favour of male goats in terms of BW and CC within breed. However, CC had an edge over Breed and Sex in BW prediction. MARS algorithm was also superior to CART, Exhaustive CHAID, CHAID and ALM models in estimating BW. The implication of the present findings is that BW can be accurately estimated from CC which saves a lot of time and labour. This can be exploited in the provision of easy and reliable information for selection of superior animals, meeting appropriate feeding and health needs including the determination of appropriate selling price for the goats.

ACKNOWLEDGEMENT

The authors would like to express their appreciation to all the farmers/marketers that facilitated the research.

Received, July 26th, 2021
Accepted November 28th, 2022

REFERENCES

ABD-ALLAH, S., H.H., ABD-EL RAHMAN, M.M., SHOUKRY, M.I., MOHAMED, F.M. SALMAN, A.A. ABEDO (2019): Some body measurements as a management tool for Shami goats raised in subtropical areas in Egypt. Bulletin of the National Research Centre, *43*, 17, https://doi.org/10.1186/s42269-019-0042-9.

AKIN, M., S.P. EYDURAN, E. EYDURAN, B.M. REED (2020): Analysis of macro nutrient related growth responses using multivariate adaptive regression splines. Plant Cell, Tissue and Organ Culture (in press), https://doi.org/10.1007/s11240-019-01763-8.

AKKOL, S. (2018): The prediction of live weight of hair goats through penalized regression methods: LASSO and adaptive LASSO. Archives of Animal Breeding, *61*: 451-458,

ALI, M., E. EYDURAN, M.M. TARIQ, C. TIRINK, F. ABBAS, M.A. BAJWA, M.H. BALOCH, NIZAMANI, A. WAHEED, M.A. AWAN, S.H. SHAH, Z. AHMAD, S. JAN (2015): Comparison of artificial neural network and decision tree algorithms used for predicting live weight at post weaning period from some biometrical characteristics in Harnai Sheep. Pakistan Journal of Zoology, *47*: 1579-1585.

ALVES, A.A.C., A.C. PINZON, R.M. DACOSTA , M.S. DASILVA, E.H.M. VIEIRA, I.B. DEMENDONÇA, S.S. VIANA, R.N.B. LÔBO (2020): Multiple regression and machine learning based methods for carcass traits and saleable meat cuts prediction using non-invasive in vivo measurements in commercial lambs. Small Ruminant Research, *171*: 49-*56*.

AMEEN, A.A., N. MIKAIL (2018): Live body weight prediction in hair goats by application of fuzzy logic. Applied Ecology and Environmental Research, *16*: 7563-7574.

BEDADA, Z.E., B.N. GILO, G.T. DEBELA (2019): Morphometric and physical characterization of Borana indigenous goats in southern Oromia, Ethiopia. Universal Journal of Agricultural Research, ***7***: 25-31.

CAMPOS, L.M., A.K. ALMEIDA, B. BIAGIOLI, K.T. RESENDE, I.A.M.A. TEIXEIRA (2017): Predicting empty body weight in growing goats: A meta-analytic approach. Small Ruminant Research, *155*: 45-50.

CANNAS, A., L.O. TEDESCHI, A.S. ATZORI, M.F. LUNESU (2019): How can nutrition models increase the production efficiency of sheep and goat operations?, Animal Frontiers, 9, 33-44.

CANUL-SOLIS, J., J.C. ANGELES-HERNANDEZ, R.A. GARCÍA-HERRERA, O.E. DEL RAZO-RODRÍGUEZ, H.A.L. RANGEL, A.T. PIÑEIRO-VAZQUEZ, F. CASANOVA-LUGO, C.A.R. NIETO, A.J. CHAY-CANUL (2020): Estimation of body weight in hair ewes using an indirect measurement method. Tropical Animal Health and Production, *52*: 2341–2347.

CELIK, S. (2019): Comparing predictive performances of tree-based data mining algorithms and MARS algorithm in the prediction of live body weight from body traits in Pakistan goats. Pakistan Journal of Zoology, *51*: 1447-1456.

CELIK, S., O. YILMAZ (2018): Prediction of body weight of Turkish Tazi dogs using data mining techniques: Classification and regression tree (CART) and multivariate adaptive regression splines (MARS). Pakistan Journal of Zoology, *50*: 575-583.

CELIK, S., E. EYDURAN, K. KARADAS, M.M. TARIQ (2017): Comparison of predictive performance of data mining algorithms in predicting body weight in Mengali rams of Pakistan. Revista Brasileira de Zootecnia, *46*: 863-872, http://dx.doi.org/10.1590/s1806-92902017001100005

CORA, E.C., A. CONDORI, J.L. CONTRERAS, J. CURASMA, A.G. CORDERO, N. VALENCIA, P.H. MAYHUA, B.A. MCGREGOR (2019): Biometric characteristics in vicuñas (*Vicugna vicugna mensalis*). Small Ruminant Research, *175*: 52-56.

CORRON, L., F. MARCHAL, S. CONDEMI, K. CHAUMOITRE, P. ADALIAN (2016): A new approach of juvenile age estimation using measurements of the ilium and multivariate adaptive regression splines (MARS) models for better age prediction. *Journal of Forensic Sci*ence, 2016, https://doi.org/ 10.1111/1556-4029.13224

DAKHLAN, A., A. QISTHON, M.D.I. HAMDANI (2021): Predicting body weight based on body measurements at different ages and sex in Saburai goat. Advances in Animal and Veterinary Sciences, *9*: 1791-1799.

DORANTES-CORONADO, E.J., G. TORRES-HERNÁNDEZ, O. HERNÁNDEZ-MENDO, R. ROJO-RUBIO (2015): Zoometric measures and their utilization in prediction of live weight of local goats in southern México. Springerplus, *4*: 695, https://doi.org/10.1186/s40064-015-1424-6.

ECK, K., E. KUNZ, C. MENDEL, G. LÜHKEN, I. MEDUGORAC (2019): Morphometric measurements in lambs as a basis for future mapping studies. Small Ruminant Research, *181*: 57-64.

EYDURAN, E., M. AKIN, S.P. EYDURAN (2019a). Application of multivariate adaptive regression splines through R Software. Nobel Academic Publishing, Ankara

EYDURAN, E. (2019b): ehaGoF: Calculates Goodness of Fit Statistics. R package version 0.1.0. https://CRAN.R-project.org/package=ehaGoF

EYDURAN, E. (2016a): The possibility of using data mining algorithms in prediction of live body weights of small ruminants. Advanced Calculation and Analysis, *1*: 18-21.

EYDURAN, E., I. KESKIN, Y.E. ERTURK, B. DAG, A. TATLIYER, C. TIRINK, R. AKSAHAN, M.M. TARIQ (2016b): Prediction of fleece weight from wool characteristics of sheep using regression tree method (CHAID algorithm). Pakistan Journal of Zoology, *48*: 957-960.

EYDURAN, E., D. ZABORSKI, A. WAHEED, S. CELIK, K. KARADAS, W. GRZESIAK (2017): Comparison of the predictive capabilities of several data mining algorithms and multiple linear regression in the prediction of body weight by means of body measurements in the indigenous Beetal goat of Pakistan. Pakistan Journal of Zoology, *49*: 273-282, http://dx.doi.org/10.17582/journal.pjz/2017.49.1.273.282

EYDURAN, E., A. WAHEED, M.M. TARIQ, F. IQBAL, S. AHMAD (2013): Predıctıon of lıveweıght from morphologıcal characterıstıcs of commercıal goat ın Pakıstan usıng factor and prıncıpal component scores ın multıple lınear regressıon. Journal of Animal and Plant Sciences, *23*: 1532-1540.

FAHIM, A., B.H.M. PATEL, V.V. RIJASNAZ (2013): Relationship of body weight with linear body measurements in Rohilkhand local goats. Indian Journal of Animal Research, *47*: 521-526.

FAOSTAT (2017): Food and Agriculture Organization of the United Nations. *FAOSTAT Database on Agriculture*. Rome, Italy: Accessed 12th August, 2019, http://faostat.fao.org/default.aspx

GARCÍA-MUÑIZ, J. G., R. RAMÍREZ-VALVERDE, R. NÚÑEZ-DOMÍNGUEZ, J.A. HIDALGO-MORENO (2019): Dataset on growth curves of Boer goats fitted by ten non-linear functions. Data in Brief, *23*: 103672.

GIZAW, S., T. GETACHEW, S. GOSHME, A. VALLE-ZÁRATE, J. VAN ARENDONK, S. KEMP, A.O. NWAI, T. DESSIE (2014): Efficiency of selection for body weight in a cooperative village breeding program of Menz sheep under smallholder farming system. Animal, *8*: 1249-1254.

GORCZYCA, M.T., H.F. MAIAMILAN, A.S. CAMPOSMAIA, K.G. GEBREMEDHIN (2018): Machine learning algorithms to predict core, skin, and hair-coat temperatures of piglets. Computers and Electronics in Agriculture, *151*: 286-294, https://doi.org/10.1016/j.compag.2018.06.028

HABIB, M.A., A. AKHTAR, A.K. FAZLUL, H. BHUIYAN, M.P. CHOUDHURY, M.F. AFROZ (2019): Biometrical relationship between body weight and body measurements of Black Bengal Goat (BBG). Current Journal of Applied Science and Technology, *35*: 1-7.

IBM CORP. RELEASED (2015): IBM SPSS Statistics for Windows, Version 542 23.0. Armonk, NY: IBM Corp.

ISAAC, J.L. (2005): Potential causes and life-history consequences of sexual size dimorphism in mammals. Mammal Review, *35*: 101–115. KASS, G.V. (1980): An exploratory technique for investigating large quantities of categorical data. Applied Statistics, *29*: 119-127.

KOC, E.K., H. BOZDOGAN (2015): Model selection in multivariate adaptive regression splines (MARS) using information complexity as the fitness function. Machine Learning, *101*: 35–58.

KOVALCHUK, I.Y., Z. MUKHITDINOVA, T. TURDIYEV, G. MADIYEVA, M. AKIN, E. EYDURAN, B.M. REED (2018): Nitrogen ions and nitrogen ion proportions impact the growth of apricot (*Prunus armeniaca*) shoot cultures. Plant Cell, Tissue and Organ Culture *133*: 263–273.

LANDAU, S.Y. (2017): Introductory Chapter: Is There a Future for Goat Pastoral Systems? In: Simões J., Gutiérrez C. (eds) Sustainable Goat Production in Adverse Environments: Vol. I. Springer, Cham, pp. 1-11.

LATIFI, M., M. RAZMKABIR (2019): Estimation of genetic trends for body weight traits in Markhoz goat at different ages. Spanish Journal of Agricultural Research, *17*: e04SC01. https://doi.org/10.5424/sjar/2019171-13608

LYAM, A. (2000): Nasarawa State. In: [Mamman, A.B., Oyebanji, J.O. and Peters, S.W. (eds)], Nigeria: A People United, a Future Assured. Survey of States, Vol. 2(2), Federal Ministry of Information, Abuja.

MILBORROW, S. (2011): Multivariate Adaptive Regression Splines: Package 'earth' derived from MDA:MARS by Trevor Hastie and Rob Tibshirani. Available at: https://cran.r-project.org/web/packages/earth/earth.pdf (Accessed on 09 February, 2019).

MILBORROW, S. (2018): *Notes on the earth package*. Available at: http://www.milbo.org/doc/earthnotes. pdf (Accessed on 09 February, 2019).

MOKOENA, K., K.M. MOLABE, M.C. SEKGOTA, T.L. TYASI (2022): Predicting body weight of Kalahari Red goats from linear body measurements using data mining algorithms. Veterinary World, *15*:1719-1726.

OKPEKU, M., A. YAKUBU, S.O. PETERS, M.O. OZOJE, C.O.N. IKEOBI, O.A ADEBAMBO, I.G. IMUMORIN (2011): Application of multivariate principal component analysis to morphological characterization of indigenous goats in Southern Nigeria. Acta Agriculturae Slovenica, *98*: 101–109.

OLATUNJI-AKIOYE, A.O., O.K. ADEYEMO (2009): Live weight and chest girth correlation in commercial sheep and goat herds in Southwestern Nigeria. International Journal of Morphology, *27*: 49-52.

OLFAZ, M., C. TIRINK, H. ÖNDER (2019): Use of CART and CHAID Algorithms in Karayaka Sheep Breeding. Kafkas Üniversitesi Veteriner Fakültesi Dergisi, *25*: 105-110.

OSENI, S.O., A. YAKUBU, A.R. AWORETAN (2017): Nigerian West African Dwarf goats. In: Sustainable goat production in adverse environments: Vol. 2. SIMÕES, JOÃO, GUTIÉRREZ, CARLOS (Eds.). Springer International Publishing AG, Gewerbestrasse 11, Switzerland, pp. 91-110.

POLAK, J., D. FRYNTA (2010): Patterns of sexual size dimorphism in cattle breeds support Rensch's rule. Evolutionary Ecology, *24*: 1255–1266.

RAD NAROUI R.M., G. KEYKHA, J. ABBASKOOHPAYEGANI, R. RAFEZI (2020): Machine learning approaches to classify melon landraces based on phenotypic traits. Genetika, *52*: 1021-1029.

RAMOS, I.O., M.P.G. DE REZENDE, P.L.S., CARNEIRO, J.C. SOUZA, J.R. SERENO, R. BOZZI, C.H.M. MALHADO (2019): Body conformation of Santa Inês, Texel and Suffolk ewes raised in the Brazilian Pantanal. Small Ruminant Research, *172*, 42-47.

RATIVA, D., B. FERNANDES, A. ROQUE (2018): Height and weight estimation from anthropometric measurements using machine learning regressions. IEEE Journal of Translational Engineering in Health and Medicine, *6*: 4400209.

R CORE TEAM (2014): R: A language and environment for statistical computing. R Foundation for Statistical Computing, 2014. Available at: http://www.Rproject.org

SONG, Y-Y., Y. LU (2015): Decision tree methods: applications for classification and prediction. Shanghai Archives of Psychiatry, *27*, 130-135.

SOWANDE, O.S., B.F. OYEWALE, O.S. IYASERE (2010): Age- and sex-dependent regression models for predicting the live weight of West African Dwarf goat from body measurements. Tropical Animal Health and Production, *42*: 969–975.

TEMOSO, O., M. COLEMAN, D. BAKER, P. MORLEY, L. BALESENG, A. MAKGEKGENENE, S. BAHTA (2017): Using path analysis to predict bodyweight from body measurements of goats and sheep of communal rangelands in Botswana. South African Journal of Animal Science, *47*: 854-863.

TYASI, T.L., A. YAKUBU, O.M. AKINSOLA, J. NGAMBI (2022): Evaluation of growth curve for South African non-descript indigenous goats using non-linear models. Acta Agriculturae Scandinavica A- Animal Science, DOI: 10.1080/09064702.2022.2112752

VALSALAN, J., T. SADAN, T. VENKETACHALAPATHY (2020): Multivariate principal component analysis to evaluate growth performances in Malabari goats of India. Tropical Animal Health and Production, *52*: 2451–2460.

VANVANHOSSOU, S.F.U., R.V.C. DIOGO, L.H. DOSSA (2018): Estimation of live bodyweight from linear body measurements and body condition score in the West African Savannah Shorthorn cattle in North-West Benin. Cogent Food & Agriculture, *4*: https://doi.org/10.1080/23311932.2018.1549767.

WORKU, A. (2019): Body weight had highest correlation coefficient with heart girth around the chest under the same farmers feeding conditions for Arsi Bale sheep. International Journal of Agricultural Science and Food Technology, *5*: 6-12.

YAKUBU, A., L. DAHLOUM, E.G. GIMBA (2019): Smallholder cattle farmers' breeding practices and trait preferences in a tropical guinea savanna agro-ecological zone. Tropical Animal Health and Production, *51*: 1497-1506.

YAKUBU, A., O.I.A., OLUREMI, E.I. EKPO (2018): Predicting heat stress index in Sasso hens using automatic linear modeling and artificial neural network. International Journal of Biometeorology, *62*: 1181-1186.

YAKUBU, A., O.I.A. OLUREMI, Z.N. IBRAHIM (2018): Modelling egg production in Sasso dual-purpose birds using linear, quadratic, artificial neural network and classification regression tree methods in the tropics. Livestock Research for Rural Development 30, Article #172. Retrieved October 1, 2018, from http://www.lrrd.org/lrrd30/10/abdul30172.html

YAKUBU, A., A.E. SALAKO, M. DE DONATO, M.I. TAKEET, S.O. PETERS, M. OKPEKU, M. WHETO, I.G. IMUMORIN (2016): Interleukin-2 (IL-2) gene polymorphism and association with heat tolerance in Nigerian goats. Small Ruminant Research, *141*: 127-134. YAKUBU, A., G.L. MOHAMMED (2012): Application of path analysis methodology in assessing the relationship between body weight and biometric traits of Red Sokoto goats in Northern Nigeria. Biotechnology in Animal Husbandry, *28*: 107-117.

YAKUBU, A. (2012): Application of regression tree methodology in predicting the body weight of Uda sheep. Scientific Papers: Animal Science and Biotechnologies, *45*, 484-490.

YAKUBU, A. (2010). Fixing multicollinearity instability in the prediction of body weight from morphometric traits of White Fulani cows. Journal of Central European Agriculture, *11*, 487-492.

YANG, H. (2013): The Case for being Automatic: Introducing the Automatic Linear Modeling (LINEAR) Procedure in SPSS Statistics. Multiple Linear Regression Viewpoints, *39*: 27-37.

YUSUF, A.I., A. ARUWAYO, I.R. MUHAMMAD (2018): Characterisation of small ruminant production systems in semi-arid urban areas of northern Nigeria. Journal of Applied Sciences and Environmental Management, *22*: 725-729.

ZAKERI, I.F., A.L. ADOLPH, M.R. PUYAU, F.A. VOHRA, N.F. BUTTE (2010): Multivariate adaptive regression splines models for the prediction of energy expenditure in children and adolescents. Journal of Applied Physiology, *108*, 128–136. ZERGAW, N., T. DESSIE, K. KEBEDE (2017): Using morphometric traits for live body weight estimation and multivariate analysis in Central Highland and Woyto-Guji goat breeds, Ethiopia. African Journal of Agricultural Research, *12*: 1326-1331.

UPOTREBA LINEARNOG MODELIRANJA, MULTIVARIJANTNE ADAPTIVNE REGRESIJE I STABLA ODLUČIVANJA U PREDVIĐANJU TELESNE TEŽINE KOD KOZA

Abdulmojeed YAKUBU[1]*, Ecevit EIDURAN[2], Senol CELIK[3], Juliana O. ISHAIA[1]

[1]Odsek za nauku o životinjama, Poljoprivredni fakultet, Državni univerzitet Nasarava, Keffi, kampus Šabu-Lafija, P.M.B. 135, Lafia, 950101, Nigerija
[2] Odeljenja za nauku o životinjama i poslovnu administraciju, Univerzitet Igdir, Igdir, Turska
[3]Odsek za nauku o životinjama, Poljoprivredni fakultet, Univerzitet Bingol, Bingol, Turska

Izvod

Upotreba robusnih algoritama regresije za bolje predviđanje telesne težine (BV) dobija sve veću pažnju. Ova studija je stoga imala za cilj predviđanje telesne mase na osnovu obima grudnog koša, rase i pola ukupno 1.012 koza. Životinje su se sastojale od 332 zrela zapadnoafrička patuljka (VAD) (197 dolara i 135 dolara), 374 crvenog sokota (RS) (216 dolara i 158 dolara) i 306 sahela (SH) (172 dolara i 134 dolara) nasumično odabranih u državi Nasarava , sever centralne Nigerije. BV predviđanje je napravljeno korišćenjem automatskog linearnog modeliranja (ALM), multivarijantnih adaptivnih regresionih splajnova (MARS), stabla klasifikacije i regresije (CART), hi-kvadrat automatske detekcije interakcije (CHAID) i iscrpnog CHAID-a. Prediktivna sposobnost svakog statističkog pristupa je merena korišćenjem kriterijuma dobrote uklapanja, tj. Pirsonovog koeficijenta korelacije (r), koeficijenta determinacije (R2), prilagođenog koeficijenta determinacije (Adj. R2), srednje kvadratne greške (RMSE), srednje vrednosti apsolutna procentualna greška (MAPE), srednja apsolutna devijacija (MAD), globalna relativna greška aproksimacije (RAE), odnos standardne devijacije (SD ratio), Akaike-ov informacioni kriterijum (AIC) i Akaike-ov informacioni kriterijum korigovan (AICc). Muškarci RS i SH koze su imale značajno ($P<0,05$) veće BV i CC u poređenju sa svojim ženkama, dok su u VAD muške koze imale značajno ($P<0,05$) veće CC ($57,88\pm0,51$ naspram $55,45\pm0,55$). Utvrđeno je da je CC osobina od najveće važnosti u predviđanju BV, kao što se i očekivalo. Među pet modela, MARS algoritam se najbolje uklapa u BV predviđanje sa r, R2, Adj. R2, SDratio, RMSE, RAE, MAPE, MAD, AIC i AICc vrednosti od 0,966, 0,933, 0,932, 0,26, 1,078, 0,045, 3,245, 0,743, 186,0 i 187, respektivno. Sadašnje informacije mogu da usmere izbor modela koji se može iskoristiti u selekciji i genetskom poboljšanju životinja, uključujući upravljanje hranom i zdravljem i marketinške svrhe, a posebno u identifikaciji standarda proučavane rase.

Primljeno 26.VII.2021.
Odobreno 28. XI. 2022.