American Journal of Animal and Veterinary Sciences

Original Research Paper

Comparison of Multivariate Adaptive Regression Splines and Classification Regression Tree for Prediction of Body Weight of Bapedi Sheep

Madumetja Cyril Mathapo, Ramaisela Barley Mthembu, Albino Joas Tsenane, Thabang Sako and Thobela Louis Tyasi

Department of Agricultural Economics and Animal Production, School of Agriculture and Environmental Sciences, University of Limpopo, Private Bag, Sovenga, Limpopo, South Africa

Article history
Received: 28-04-2024
Revised: 13-05-2024
Accepted: 14-05-2024

Corresponding Author:
Madumetja Cyril Mathapo
Department of Agricultural Economics and Animal Production, School of Agriculture and Environmental Sciences, University of Limpopo, *Private Bag,* Sovenga, Limpopo, South Africa
Email: madumetja.mathapo@ul.ac.za

Abstract: The study aimed to compare the performance of multiple adaptive regression splines and classification regression trees for the prediction of the body weight of Bapedi sheep. A total of 100 Bapedi sheep aged between one and five years old of different sexes were employed. The study measured the following: Body Length (BL), Withers Height (WH), Heart Girth (HG), Rump Height (RH), Body Weight (BW) and Sternum Height (SH). The model's performances were evaluated using goodness of fit criteria while the association between body measures and BW was discovered using a correlation matrix. Multivariate Adaptive Regression Splines (MARS) and Classification and Regression Tree (CART) established the model for the prediction of BW. The findings indicated that CART performed well. Correlation matrix results indicated that BW had positive statistical significance ($p<0.05$) with SH ($r = 0.53$) and BL (0.47) and a statistically significant relationship ($p<0.01$) with HG ($r = 0.89$), WH ($r = 0.74$) and RH ($r = 0.64$). CART model indicated that HG, BL, and BL could be used to predict BW while MARS indicated that HG, the interaction of AGE and HG, and the interaction of AGE, BL, and HG play a role in the prediction of BW. The CART model appears to be the most effective model for predicting BW based on goodness of fit results. The correlation results imply that HG can be applied to enhance the BW of Bapedi sheep. CART and MARS model results suggest that HG and the interaction of AGE and HG are the best explanatory variables of BW.

Keywords: CART MARS, Body Measurements, Correlations, Goodness of Fit Criterion

Introduction

Sheep a small ruminant that has served a variety of functions in most civilizations and in rural development (Celik *et al*., 2017). The Bapedi sheep is one of the Indigenous breeds found in South Africa and originated from the hot temperatures that receive low moderate rainfall known as Limpopo province (Maqhashu *et al*., 2020). This breed does not require a lot from farmers due to its adaptability to harsh environmental conditions and ability to survive on lack of good quality feed and tolerance to diseases (Kunene *et al*., 2014). Bapedi sheep are smaller fat-tailed sheep with white, brown, and red colors and are known to be kept for meat at a rural level (Duncanson, 2012). Bapedi sheep play an important role in the livelihoods of smallholder farmers since they fight poverty and food insecurity. (Ngcobo *et al*., 2022). However, the breed has been forsaken and turned to face the risk of extinction (Maqhashu *et al*., 2020). Therefore, the phenotypic characterization of sheep (Ndiaye *et al*., 2019) and the development of a breeding program (Kebede and Usman, 2023) can enhance and sustain the production of livestock. However, in phenotypic characterization and breeding programs, it is important to know the association between body weight and the connected traits since are correlated genetically (Celik *et al*., 2017). Understanding about body weight of sheep can assist in knowing the performance of the animal, making proper feeding and selecting the best animals for breeding (Abbas *et al*., 2021). However, at a rural level, farmers are unable to estimate the BW of animals since there are no scales (Mahmud *et al*., 2014). Body weight has been predicted from body measurements through multiple regression (Birteeb and Ozoje, 2012), stepwise regression (Kumar *et al*., 2018), and simple regression (Ahmed *et al*., 2017). This regression model does not give a precise estimation of body weight due to the

© 2024 Madumetja Cyril Mathapo, Ramaisela Barley Mthembu, Albino Joas Tsenane, Thabang Sako and Thobela Louis Tyasi. This open-access article is distributed under a Creative Commons Attribution (CC-BY) 4.0 license.

multicollinearity problem that occurs when there is a high correlation among the body measurements (explanatory variables) (Iqbal *et al*., 2023). According to Vázquez-Martínez *et al*. (2023), these algorithms (MARS and CART) are not affected by multicollinearity problems and can provide one with accurate estimations.

According to our knowledge, there is limited information on the prediction of BW from some morphometric traits of Bapedi sheep using CART and MARS. Thus, the study's goal is to (1) Determine the association between body measurements of Bapedi sheep and body weight. (2) Establish a model to utilize for the prediction of body weight from some body measurements using CART and MARS. This study will come up with a trait that can be used to select animals for breeding and prediction of body weight in the absence of scales.

Materials and Methods

Study Area

The study was carried out at Lepel Lane farm found at Ga-Nkwana village, under Fetakgomo local municipality. Ga-Nkwana settlement (24° 24' 22" S 29° 47' 17" E) is marked by semi-arid climates with very unpredictable precipitation patterns and the village receives a typical yearly precipitation amount of 551 mm, with a highest monthly rainfall of 267 mm in the summer months of December/January/February. The average highest daylight temperature is 32°C in January, with the mean minimum temperature being 6°C in July (Radingoana *et al*., 2019). Sekhukhune districts consist of four vegetation types viz: Plains bushveld, mountain bushveld, montane grassland, and summit sourveld (Moela, 2014).

Experimental Animal and Management

Animals of different sexes aged between one and five years old were used in the study. The animals were released in the morning for grazing and were recalled later into the kraals where they were provided with water. The animals were regularly inspected and dipped for external parasites. Sick and pregnant animals were excluded from the study.

Sampling Procedure

A census research approach was used in this study. This approach involved gathering data from each member of the statistical population. Therefore, a total of 100 Bapedi sheep were sampled from Lepel Lane farm.

Data Collection

The live body weights of Bapedi sheep were determined using a digital scale calibrated in kilograms while morphometric traits were measured through tailor measuring tape following the procedure of FAO (2012). During the measurements, the animals were restrained in a standing still manner where all four legs were straight on the ground. The measurements were taken by one person to avoid individual variation. The sex of the animals was determined visually while age was determined through the dental chart.

Statistical Analysis

Analysis was run by R studio software. Objective one was achieved through a correlation matrix while (CART) and MARS were used to achieve objective two. Goodness of fit was used to check the performance accuracy of the model used. The predictive performance of the models was tested following the procedure of Celik (2019).

Results

Figure 1 demonstrates the correlation between body weight and a number of body measurement characteristics. The results showed a correlation ($p<0.01$) between BW and HG (0.89), WH (0.74), RH (0.64), and a mid-correlation ($p<0.05$) with SH (0.53) and BL (0.47). The findings also showed body measurements used in the study had a mid-significant ($p<0.5$) and high-significant correlation ($p<0.01$) among them.

Table 1 shows the findings of the MARS model. The first term of the model was the intercept. The model showed an intercept of 27.72, while the second term and first basic function was HG with a cut point of <66 cm and a -0.80-regression coefficient. The second period and the second fundamental role was HG with a cut point of >66 cm with a regression coefficient of 1.03. The second term and the third and fourth basic functions were the interaction of AGE3* h (HG>66 cm) and AGE3*BL* h (HG>66 cm) with regression coefficients affecting the model positively and negatively with values of 5.81 and -0.08, respectively.

Table 1: MARS model

Basic Functions (BF)		Coefficients
Intercept		27.72
h (66-HG)	BF1	-0.80
h (HG-66)	BF2	1.03
AGE3* h (HG-66)	BF3	5.81
AGE3* BL* h (HG-66)	BF4	-0.08

HG: Heart Girth; BL: Body Length

Figure 2 shows a tree diagram obtained from the CART model. The root node of the tree diagram had a typical BW of 36 kg indicating 100% among the flock. Using HG and 73 cm, the two subgroups were formed out of the base node. Within the initial subgroup, if HG is less than 73 cm, the average BW was determined to be 29 kg, accounting for 44% of the flock; if HG is greater than 73 cm, the typical BW was 42 kg, accounting for 56% of the flock. By using HG, the tree was split into two subgroups of 69 and 81 cm. The average BW was 29 kg, or 19% of the flock, where the HG was less than 69 cm, and 32 kg, or 25% of the flock, more than 67 cm separated the HG. When the HG was less than 81 cm, the average BW was 40 kg, while the average BW was 50 kg when the HG was greater than 81 cm. The tree was divided into two subgroups via BL with 52 and SH 50 cm. Where 52<BL>52 the average BW was 20 and 26 kg representing 6 and 14% of the flock, respectively. When 50≤SH≥50 cm the average BW was 45 kg and 54 kg representing 7% and 8% of the flock, respectively. The tree diagram had 6 terminal nodes.

Table 2 shows the predictive performance of the CART and mart model. From both the CART and MARS model it was shown that the values from the train set outperformed values from the test set. When CART and MARS outcomes were compared during training, they were nearly identical. The results indicated that the best model was CART and was obtained from the training data set. The training data set of the CART model had the highest coefficients of determination (0.87), adjusted coefficient of determination (0.87) and Pearson correlation (0.93) and lowest Akaike information criteria (170.28), Root mean square error (3.26) and relative approximate error (0.01).

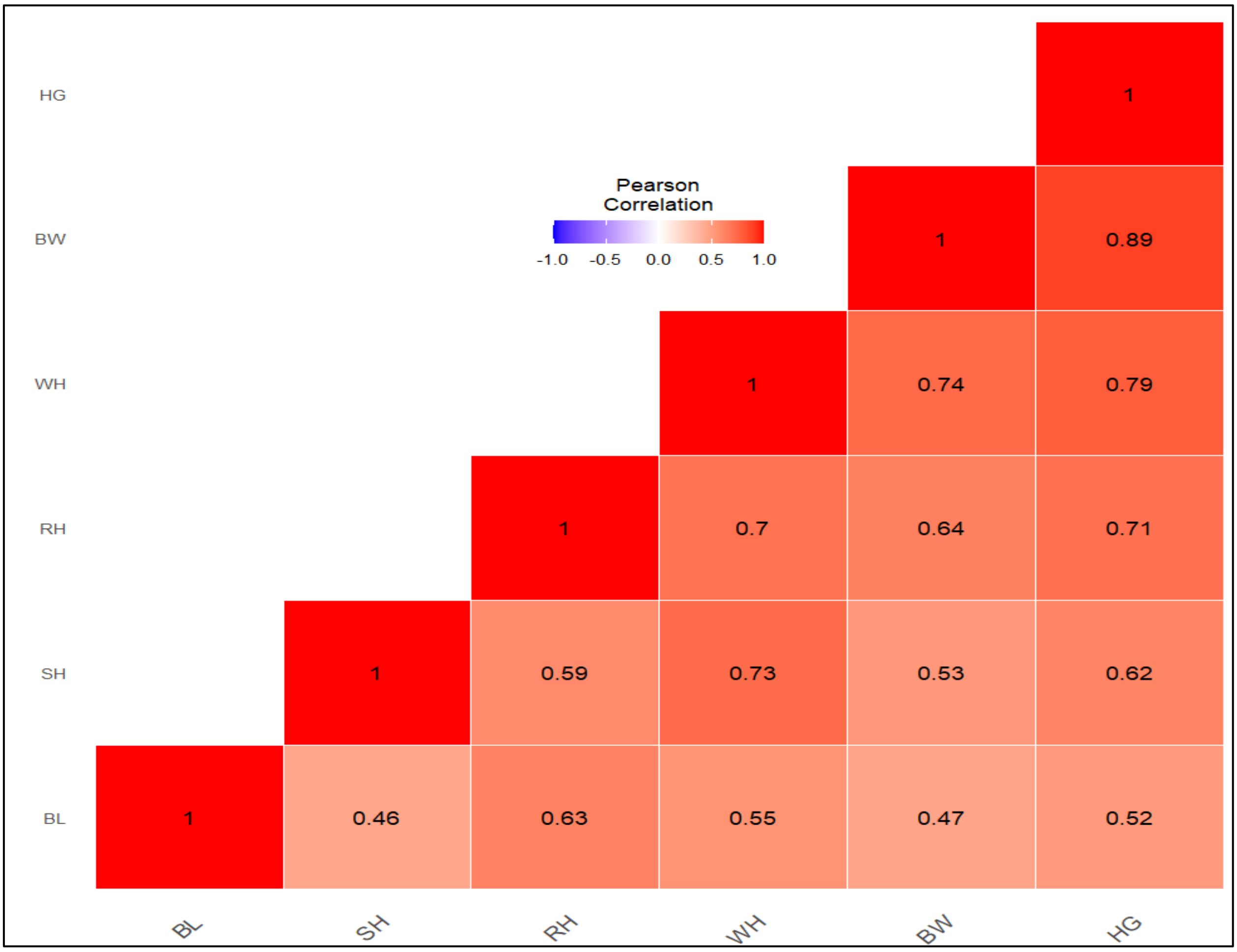

Fig. 1: Heat map displaying an association between body weight and morphometric traits of Bapedi Sheep. Heat map color distribution that represents the magnitude and direction of the association between the dependent and independent variables. Where: Red color means high correlation (p<0.01), white color means mid-correlation (p<0.05), and blue color means low correlation with negative direction

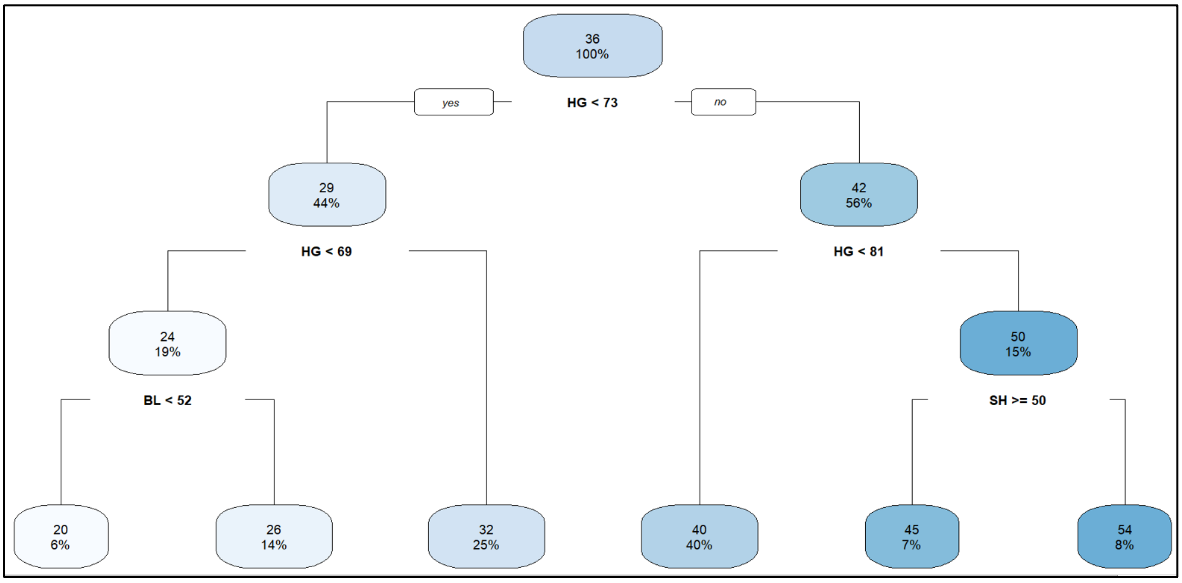

Fig. 2: Regression decision tree diagram for body weight in Bapedi sheep using the CART algorithm

Table 2: The goodness of fit criteria results

	CART		MARS	
Criterion	Train	Test	Train	Test
Root Mean Square Error (RMSE)	003.26	5.33	3.97	11.14
Relative Root Mean Square Error (rRMSE)	008.98	14.64	10.92	11.14
Standard Deviation Ratio (SDr)	000.36	0.59	0.44	0.46
Coefficient of Variation (CV)	009.04	13.70	10.99	10.63
Pearson Correlation (PC)	000.93	0.83	0.90	0.89
Relative Approximate Error (RAE)	000.01	0.02	0.01	0.01
Mean Relative Approximate Error (MRAE)	000.10	0.03	0.01	0.02
Mean Absolute Percentage Error (MAPE)	006.27	13.38	7.43	7.39
Mean Absolute Deviation (MAD)	002.32	4.15	2.80	2.90
Coefficient of Determination (Rsq)	000.87	0.59	0.81	0.76
Adjusted Coefficient of Determination (AdjRsq)	000.87	0.59	0.80	0.73
Akaike Information Criterion (AIC)	170.28	93.67	204.41	84.40

Discussion

The use of data mining algorithms can help researchers guide farmers to determine which characteristics to use in order to forecast the weight of their sheep in case there is no scale, to track the growth of their animals, and run proper animal management (Vázquez-Martínez *et al*., 2023). The objective of the study was to check the association between morphometric traits using a correlation matrix. Results indicated that withers height, heart girth, and rump height were statistically connected with body weight, while on the other hand, it had a statistical correlation with body length and sternum height. Results were in accordance with the findings of Ali Rather *et al*. (2020) who discovered a statistically significant relationship between Kashmir merino sheep's body length and weight. The current study results were in agreement with the findings of Kumar *et al*. (2018); Sam *et al*. (2023), who discovered an extremely substantial statistical relationship between the withers height, heart girth, and body weight of West African dwarf sheep and Harnail sheep. However, Saeed *et al*. (2023), reported a non-significant correlation between body weight and all the studied traits of interest in Holstein heifer. Variation in species could contribute to the difference, in parameters that were studied.

Therefore, a body weight prediction model using morphological features was established by the study using the MARS and CART algorithms. Multivariate adaptive regression splines revealed that heart girth, interaction of age and heart girth, and interaction of age, body length, and heart girth can be used to predict the body weight of Bapedi sheep. The findings indicate that the interaction of age (3 years) and heart girth greater than sixty-six centimeters were the best indicators of body mass. As per the findings of Faraz *et al*. (2021), age five was the most effective explanatory factor in Thalli sheep. Similarly, Rashijane *et al*. (2023) reported heart girth greater than a hundred centimeters in savanna goats as the greatest predictor of body weight. Variations in these results may be caused by the various species used and the surrounding circumstances. Cart model findings indicate that heart girth, body length, and sternum height are suitable traits for the estimation of the body weight of Bapedi in the absence of a weighing scale. The results showed that the greatest explanatory variable for body weight prediction is heart girth. Tirink *et al.* (2023), reported similar results where the classification and regression tree model showed heart girth as the greatest predictor of body weight in rams. Olfaz *et al.* (2019), however, presented data from the current investigation that were different, where varied conclusions from the present investigation, wherein the CART model showed that birth type was the explanatory factor for the prediction of Karayaka sheep.

The results of Oyebanjo *et al*. (2023) where CART was used to predict the body weight of two goat breeds were consistent in the estimation of the weight of Red Sokoto and in contrast in the prediction of Sahel where body length was the best predictor of body weight. The study further looked at the performance of the CART and MARS models using goodness of fit criteria. The findings on the performance of the models indicate that the CART model performed well, where it showed greater determination coefficient and adjusted determination coefficient, Pearson correlation, lower standard ratio, relative approximate error, mean relative approximate error, and root mean square error. Tirink *et al*. (2023), reported similar results where the CART model was found as the most accurate model for Romane sheep breed body weight estimation. Celik (2019) reported different results where MARS was found to be the best predictor of the body weight of Pakistan goats. Celik *et al*. (2017) results were in accordance with the results of the study, where the CART model performed the best when estimating Pakistan Mangali ram's body weight. However, Hlokoe *et al*. (2022), reported that MARS was found to be the most accurate model for estimating the body weight of Nguni cows from the University of Limpopo. The difference might be the variation of species used and the geographic area.

Conclusion

The study concludes that there Bapedi sheep body dimensions (sternum height, body length, withers height, rump height, and heart girth) and body weight have a statistically significant positive relationship. The best variable to explain the body weight of Bapedi sheep is heart girth. Classification and regression tree indicates that heart girth plays a crucial role in the prediction of body weight of Bapedi sheep while Multivariate adaptive regression splines indicate that interaction of age (3 years) and heart girth greater than sixty-six centimeters plays a role in the body weight. The study further concluded that the classification and regression model performed well in coming up with the traits that can be employed to predict the Bapedi sheep's body weight. This information can help farmers at a rural level to select their sheep for breeding. The information will further be used to develop a weighing belt that can be used by the farmers in the absence of weighing scales to estimate the weight of their sheep.

Acknowledgment

We are grateful to the Lepallane farmer for granting us permission to perform this study using their animals.

Funding Information

The authors have not received any financial support or funding to report.

Author's Contributions

Madumetja Cyril Mathapo: Designed the study read and approved the final manuscript.

Ramaisela Barley Mthembu, Albino Joas Tsenane, and Thabang Sako: Collected, analyzed data, and wrote the manuscript.

Thobela Louis Tyasi: Read, improved, and approved the final manuscript**.**

Ethics

This study was approved by the University of Limpopo's Animal Research Ethics Committee (AREC), which also provided a certificate bearing the AREC number: AREC/42/2023:UG.

The Animal Research Ethics Committee (AREC) of the University of Limpopo approved this study and issued out the certificate with the number: AREC/42/2023:UG.

Conflict of Interest

The authors declare no competing interest.

References

Abbas, A., Ullah, M. A., & Waheed, A. (2021). Body weight prediction using different data mining algorithms in Thalli sheep: A comparative study. *Veterinary World*, *14*(9), 2332–2338. https://doi.org/10.14202/vetworld.2021.2332-2338

Ahmed, S., Omer, S., & MElzaki, R. (2017). Prediction of Live Body weight of Sudanese Goat Kids from Body Size Measurements. *International Journal of Livestock Research*, *7*, 72–79. https://doi.org/10.5455/ijlr.20170102113257

Ali Rather, M., Bashir, I., Hamdani, A., Nabi Khan, N., A. Ahangar, S., & Nazki, M. (2020). Prediction of Body Weight from Linear Body Measurements in Kashmir Merino Sheep. *Advances in Animal and Veterinary Sciences*, *9*(2), 189–193. https://doi.org/10.17582/journal.aavs/2021/9.2.189.193

Birteeb, P. T., & Ozoje, M. O. (2012). Prediction of live body weight from linear body measurements of west African long-legged and west African dwarf sheep in northern Ghana. *Online Journal of Animal and Feed Research*, *2*(5), 427–434. https://www.researchgate.net/publication/257305181

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*(4), 1447–1456. https://doi.org/10.17582/journal.pjz/2019.51.4.1447.1456

Celik, S., Eyduran, E., Karadas, K., & Tariq, M. M. (2017). Comparison of predictive performance of data mining algorithms in predicting body weight in Mengali rams of Pakistan. *Revista Brasileira de Zootecnia*, *46*(11), 863–872. https://doi.org/10.1590/s1806-92902017001100005

Duncanson, G. R. (2012). *Veterinary Treatment of Sheep and Goats* (1st Ed.). CABI, 2012. ISBN-10: 9781780640051.

FAO. (2012). Phenotypic characterization of animal genetic resources. Animal Production and Health Guidelines No. 11. *Animal Genetic Resources*, *50*, 110. https://doi.org/10.1017/s2078633612000239

Faraz, A., Tirink, C., Eyduran, E., Waheed, A., Tauqir, N. A., Nabeel, M. S., & Tariq, M. M. (2021). Prediction of live body weight based on body measurements in Thalli sheep under tropical conditions of Pakistan using CART and MARS. *Tropical Animal Health and Production*, *53*(2), 1–12. https://doi.org/10.1007/s11250-021-02748-6

Hlokoe, V. R., Mokoena, K., & Tyasi, T. L. (2022). Using multivariate adaptive regression splines and classification and regression tree data mining algorithms to predict body weight of Nguni cows. *Journal of Applied Animal Research*, *50*(1), 534–539. https://doi.org/10.1080/09712119.2022.2110498

Iqbal, F., Raziq, A., Tirink, C., Fatih, A., & Yaqoob, M. (2023). Using the artificial bee colony technique to optimize machine learning algorithms in estimating the mature weight of camels. *Tropical Animal Health and Production*, *55*(2), 86. https://doi.org/10.1007/s11250-023-03501-x

Kebede, K., & Usman, A. (2023). Valuing the Relative Importance of Goat Farmers Trait Preferences: Implication for Designing Community Based Breeding Program. *Journal of Biomedical Research & Environmental Sciences*, *4*(10), 1383–1393. https://doi.org/10.37871/jbres1808

Kumar, S., Dahiya, S. P., Dahiya, Z. S., & Patil, C. S. (2018). Prediction of body weight from linear body measurements in sheep. *Indian Journal of Animal Research*, *52*(9), 1263–1266. https://doi.org/10.18805/ijar.b-3360

Kunene, N. W., Ceccobelli, S., Lorenzo, P. D., Hlophe, S. R., Bezuidenhout, C. C., & Lasagna, E. (2014). Genetic diversity in four populations of Nguni (Zulu) sheep assessed by microsatellite analysis. *Italian Journal of Animal Science*, *13*(1), 3083. https://doi.org/10.4081/ijas.2014.3083

Maqhashu, A., Mapholi, N. O., O'Neill, H. A., Nephawe, K. A., Ramukhithi, F. V., Sebei, J. P., Nxumalo, K. S., & Nedambale, T. L. (2020). Assessment of genetic variation in Bapedi sheep using microsatellite markers. *South African Journal of Animal Science*, *50*(2), 318–324. https://doi.org/10.4314/sajas.v50i2.15

Moela, A. K. (2014). *Assessment of the relationship between body weight and body measurement in indigenous goats using path analysis* [University of Limpopo]. https://doi.org/http://hdl.handle.net/10386/1317

Ndiaye, B., Diouf, M. N., Ciss, M., Wane, M., Diop, M., & Sembène, M. (2019). Phenotypic characterization of Senegalese Peul-peul sheep. *International Journal of Advanced Research*, *7*(2), 1021–1027. https://doi.org/10.21474/ijar01/8575

Ngcobo, J. N., Nedambale, T. L., Nephawe, K. A., Mpofu, T. J., Chokoe, T. C., & Ramukhithi, F. V. (2022). An Update on South African Indigenous Sheep Breeds' Extinction Status and Difficulties during Conservation Attempts: A Review. *Diversity*, *14*(7), 516. https://doi.org/10.3390/d14070516

Mahmud, M., Shaba, P., Abdulsalam, W., Yisa, H., Gana, J., Ndagi, S., & Ndagimba, R. (2014). Live body weight estimation using cannon bone length and other body linear measurements in Nigerian breeds of sheep. *Journal of Advanced Veterinary and Animal Research*, *1*(4), 169–176. https://doi.org/10.5455/javar.2014.a29

Olfaz, M., Tirink, C., & Önder, H. (2019). CART ve CHAID Algoritmalarının Karayaka Koyun Islahında Kullanımı. *Kafkas Üniversitesi Veteriner Fakültesi Dergisi*, *25*(1), 105–110. https://doi.org/10.9775/kvfd.2018.20388

Oyebanjo, M., Coker, O., & Osaiyuwu, O. (2023). Predicting the body weight of indigenous goat breeds from morphological measurements using the classification and regression tree (CART) data mining algorithm. *Biotechnology in Animal Husbandry*, *39*(1), 33–49. https://doi.org/10.2298/bah2301033o

Radingoana, M. P., Dube, T., Mollel, M. H. N., & Letsoalo, J. M. (2019). Perceptions on greywater reuse for home gardening activities in two rural villages of Fetakgomo Local Municipality, South Africa. *Physics and Chemistry of the Earth, Parts A/B/C*, *112*, 21–27. https://doi.org/10.1016/j.pce.2019.02.009

Rashijane, L. T., Mokoena, K., & Tyasi, T. L. (2023). Using Multivariate Adaptive Regression Splines to Estimate the Body Weight of Savanna Goats. *Animals*, *13*(7), 1146. https://doi.org/10.3390/ani13071146

Sam, I. M., Usoro, O. O., & Udo, A. F. (2023). Relationship of linear body measurements with live body weight in West African Dwarf sheep raised in Tropical Rainforest of Nigeria. *Journal of Animal Science and Veterinary Medicine*, *8*(2), 46–51. https://doi.org/10.31248/jasvm2023.365

Saeed, O. A., Alnori, H. M., Essa, A. H., Hameed, N. A., Shareef, M. A., Hamza, M. A., Abdulghafoor, R. T., & Attallah, O. K. (2023). Physiological Correlates of Holstein Heifer Body Weight: Implications for Management. *IOP Conference Series: Earth and Environmental Science*, *1252*(1), 012122. https://doi.org/10.1088/1755-1315/1252/1/012122

Tırınk, C., Piwczyński, D., Kolenda, M., & Önder, H. (2023). Estimation of Body Weight Based on Biometric Measurements by Using Random Forest Regression, Support Vector Regression and CART Algorithms. *Animals*, *13*(5), 798. https://doi.org/10.3390/ani13050798

Vázquez-Martínez, I., Tırınk, C., Salazar-Cuytun, R., Mezo-Solis, J. A., Garcia Herrera, R. A., Orzuna-Orzuna, J. F., & Chay-Canul, A. J. (2023). Predicting body weight through biometric measurements in growing hair sheep using data mining and machine learning algorithms. *Tropical Animal Health and Production*, *55*(5), 307. https://doi.org/10.1007/s11250-023-03717-x