Pakistan J. Zool., vol. 55(1), pp 407-411, 2023 DOI: https://dx.doi.org/10.17582/journal.pjz/20210126080128

Prediction of Live Body Weight for Thalli Sheep Using Chi-Square Automatic Interaction Detector and Multiple Linear Regression: A Comparative Study

Ansar Abbas[1]*, Muhammad Aman Ullah[2], Abdul Waheed[3] and Muhammad Asif[4]

[1]*Govt. Degree College for Boys Makhdoom Rashid, Multan, Pakistan*
[2]*Department of Statistics, Bahauddin Zakariya University, Multan, Pakistan*
[3]*Department of Livestock and Poultry Production, Bahauddin Zakariya University, Multan*
[4]*Govt. Degree College for boys Qadir Pur Raan, Multan, Pakistan*

Article Information
Received 26 January 2021
Revised 03 July 2021
Accepted 10 August 2021
Available online 08 March 2022 (early access)
Published 10 November 2022

Authors' Contribution
AA study design, data handling, formal analysis, methodology, drafted the study. MAU supervision, visualization, results interpretation. AW formal analysis, supervision, visualization, co-drafted the manuscript. MA methodology, software use and interpretation.

Key words
Body measurements, CHAID, MLR, Thalli sheep, Live body weight prediction

ABSTRACT

The aim of the present study was to compare the predictive performances of Chi-square automatic interaction detector (CHAID) and multiple linear regression (MLR) implemented in the prediction of live body weight (BW) from several body measurements such as withers height (WH), body length (BL), head length (HL), rump length (RL), head width (HW), rump width (RW), barrel depth (BD) and heart girth (HG). A total of 271 female Thalli sheep of southern Punjab were included in the study. We used different goodness of fit criteria such as coefficient of determination (R^2 %), adjusted coefficient of determination (Adj-R^2 %), standard deviation ratio (SD ratio), root mean square error (RMSE), relative approximation error (RAE), mean absolute deviation (MAD), coefficient of variation (CV %), Akaike information criterion (AIC) and Pearson correlation coefficient (r) for comparing the predictive performance of CHAID and MLR. To improve the predictive ability, minimum number of parent (20) and child nodes (10) were set for CHAID algorithm. The empirical results revealed that HL was the most significant predictor of live BW of Thalli sheep by using both methods. However, the comparative analysis shows that the CHAID algorithm performed well as compared to the MLR (CHAID *vs.* MLR: R^2 % = 79.38 *vs.* 73.27; Adj-R^2 %= 78.70 *vs.* 72.50; SD ratio=0.45 *vs.* 0.51; RMSE= 3.15 *vs.* 3.58; RAE= 0.12 vs. 0.14; MAD= 2.41 *vs.* 2.74; CV%=13.28 *vs.* 15.11; AIC =638.50 *vs.* 708.28; r= 0.89 *vs.* 0.85). Consequently, it is hoped that the results of the study on the morphological characterization of Thalli sheep might be a good reference for next sheep breeding studies.

INTRODUCTION

In recent years, small ruminants play a momentous role in developing economy of a country. Goat and sheep are well adapted to life and people are being farming these ruminants for multifactorial benefits such as, to fulfill the demands of meat, milk and wool. Different types of goat and sheep breeds are found in Pakistan. Thalli are one of the sheep breeds with massive potential in order to get the meat, milk, and wool. The main living places of Thalli breed is the desert regions of Muzaffargarh, Multan, Mianwalli, Jhang and Sargodha. One of the adaptive characteristics of Thalli breed is the remarkable tolerance of the temperature up to 50°; however, at this temperature other animals will suffer. The basic distinguishing features of the Thalli breed are relatively small head with pendulous ears. In the context of coloring, the black or brown head color is the exclusive characteristic of this animal. However, the other strains of this breed large heads with small ears. These two distinguish features make it more adaptive than other breeds. Medium sized body, romen nose and small tail are the other important characteristics of this breed. Genetically, the meat and wool production of Thalli is quite good as compared to other sheep breeds of Pakistan (Khan *et al*., 2003).

Predicting live body weight (BW) is an important topic to find out a proper feed amount, drug dose, and marketing for animal under rural condition without weighing instruments (Eyduran *et al*., 2013). In the literature,

* Corresponding author: ansarashri@gmail.com
0030-9923/2023/0001-407 $ 9.00/0

various reports found a great figure for the estimation of BW by using different body dimensions of different sheep breeds. Body dimensions are used to estimate BW quite well in the condition where weighbridges are not available (Berge, 1977; Buvanendran *et al*., 1980; Goonerwardene and Sahaayuraban, 1983). The predictive quality of the establishment in small ruminants is improved through fair and consecutive statistical techniques. Different researchers used various statistical techniques such as canonical correlation analysis (Yapark *et al*., 2008), path analysis (Yakubu, 2009), correlation analysis (Khan *et al*., 2006), use of factor scores in multiple regression (Tariq *et al*., 2012), simple regression (Chitra *et al*., 2012) and multiple linear regression (MLR) (Seifemichael., 2014) for BW prediction. By using the traditional statistical methods, the reliability of these methods can be harmfully affected by violation of the distributional assumptions, and it can lead to biased parameter estimation (Ruhil *et al*., 2013).

Various studies have been carried out to reduce the problem of multicollinearity for prediction the BW from body parameters by means of Chi-square automatic interaction detector (CHAID) (Khan *et al*., 2014; Ali *et al*., 2015; Akin *et al*., 2017; Eyduran *et al*., 2017; Arsalan *et al*., 2020); Classification and regression tree diagram (CART) (Kovalchuk *et al*., 2017); Artificial neural network (ANN) (Behzadi and Aslaminejad, 2010; Ali *et al*., 2015; Eyduran *et al*., 2017; Celik *et al*., 2017). Eyduran *et al*. (2017) predicted the BW of Pakistani Beetal goat based on six different traits (head girth, neck length, diagonal body length, belly sprung, shank circumference, and rump height) through CART, CHAID, ANN and MLR methods. In addition, Yakubu (2012) predicted BW through body parameters in UDA rams.

In Pakistan, numerous studies on BW prediction for Mengali, Balochi, Harnai, Beverigh and Rakhshani sheep are available (Mohammad *et al*., 2012; Khan *et al*., 2014; Ali *et al*., 2015). However, studies to predict the live BW for Thalli sheep found in Southern Punjab, Pakistan are scarce and are currently in need. This fact motivated us to carry out this study. The primary aim of this paper was to compare the predictive performances of CHAID and MLR for the prediction of live BW among Thalli sheep.

MATERIALS AND METHODS

The sample of 271 female sheep was obtained from different livestock experiment station (LES) of district Muzaffargarh and Multan. Random sampling approach was used for sheep selection. We considered only those healthy sheep who did not take any medication and physical disability. The required information was taken through self-administered questionnaire, comprised of age and morphologically measurements. Nine morphometrics traits i.e., body weight, withers height (WH), body length (BL), head length (HL), rump length (RL), head width (HW), rump width (RW), barrel depth (BD) and heart girth (HG) were measured on each sheep and were taken in standing position according to standard procedures. Data collection activity was made by the same person to avoid the between individual variation. Initially, descriptive analysis i.e., mean along with standard deviation (SD) and percentage co-efficient of variation (C.V %) of each quantitative variable were reported in Table I.

Table I. Descriptive statistics for body weight (kg) and body measurements (cm) of Thali sheep.

Trait	Mean ± SD	CV (%)
Body weight	23.75 ± 6.94	29.23
Withers height	63.95 ± 7.66	11.98
Body length	63.76 ± 9.10	14.28
Head length	24.64 ± 4.30	17.47
Rump length	12.93 ± 2.54	19.65
Head width	9.53 ± 1.63	17.12
Rump width	18.17 ± 4.36	23.99
Barrel depth	41.02 ± 6.28	15.31
Heart girth	68.18 ± 9.58	14.05

SD, Standard deviation; CV, Coefficient of variation

Primarily, a CHAID method was used for the prediction of live BW of Thalli sheep. This algorithm usually used to categorize those subsets of predictors that best depict the dependent variable. A ten-fold cross validation criterion was used in the study. The basic objective of CHAID technique is to minimize variation within nodes to construct homogenous subgroups in the optimal regression tree diagram with significant predictors (Kass, 1980). In this study, minimum sheep numbers for parent and child nodes were set at 20 and 10 and p-value for splitting equal to 0.05. Moreover, the Bonferroni adjustment was utilized to correct for the p-values of the best predictor at each split in the CHAID algorithm. All the statistical analysis was performed by using SPSS 23.0 software program.

We also predict the live BW of Thalli sheep using MLR analysis. The MLR is a commonly used method for developing the relationship between a dependent variable and a set of predictors. The mathematical form of multiple linear regressions is as follows:

$$Live\ BW = \beta_0 + \sum_{i=1}^{8} \beta_i X_i + \varepsilon$$

Where, Live BW is a dependent variable, β_0 is an intercept, β_i is the ith parameter, X_i is the ith independent variable, and ε is a random error. To compare the predictive performance of CHAID and MLR, different model evaluation criteria i.e., co-efficient of determination (R^2), adjusted co-efficient of determination (adj-R^2), Pearson correlation coefficient (r), co-efficient of variation (CV), standard deviation ratio (SD ratio), relative approximation error (RAE), root mean square error (RMSE), mean absolute deviation (MAD) and akaike information criterion (AIC) were used. The details of mathematical equation and other relevant statistical measures can be seen in earlier published studies (Akaike, 1973; Sugiura, 1978; Salehi *et al*., 1998; Willmott and Matsuura, 2005; Takma *et al*., 2012; Eyduran *et al*., 2017).

RESULTS AND DISCUSSION

Sheep production has played a major role in animal production, particularly in rural development. One of the main targets of the breeders in sheep reproduction is to find out the relationship between body weight and the related traits. For predicting live BW of any sheep breed, linear body measurements can be used as indirect selection criteria. In this study, we also planned to predict BW of female Thalli sheep from linear body measurements.

In this study, predictive performance of CHAID and MLR used for predicting BW by means of some body measurements in sheep has been evaluated comparatively. Table II shows the performance of the CHAID and MLR methods to predict BW. The estimated value of r, R^2 and Adj-R^2 were significantly higher ($P<0.05$) for the CHAID algorithm compared with that for MLR. On the other hand, SD ratio, RMSE, RAE, MAD, CV% and AIC were lowest for the CHAID algorithm which indicates the better predictive capabilities in comparison with MLR. The worst performance in the current research work was recorded for the MLR method. A study by Khorshidi-Jalali *et al*. (2019) also obtained low R^2 (67%) for predicting body weight in Raini Cashmere goat by using the MLR method. In line with our results, some earlier reports highlighted the biological advantage of CHAID algorithm in BW prediction. e.g., Celik and Yilmaz (2017) found that CHAID was the ideal tree-based algorithm in the estimation of BW trait. Their results for different goodness-of-fit criteria were, r= 0.84, R^2= 71.57, SD ratio = 0.53. Olfaz *et al*. (2019) published a study including 366 karayaka sheep in which they suggested that CHAID algorithm was the more useful method for body weight prediction as compared to the classification and regression tree (CART) method. We also compared the statistical results of different criterion with earlier reports. The present R^2 estimated for CHAID algorithm was highest while, Mohammad *et al*. (2012) estimated a lower R^2 value of 0.72 for CHAID algorithm in the body weight prediction through withers height, chest girth, body length, and breed in indigenous Pakistani sheep. However, in Harnai breed, Ali *et al*. (2015) presented better goodness-of-fit results (R^2= 83.77, r=0.91, SD ratio= 0.40, RMSE= 1.50) for CHAID algorithm compared with the corresponding estimates in the present study. Another study used the exhaustive CHAID algorithm for BW estimation of Harnai sheep and the reported results (R^2 = 0.84) were also better than our results (Khan *et al*., 2014). Aksahan (2015) apply CHAID algorithm to get morphological measurements affecting final live weight (FLW) at fattening period of 103 young (Holstein, Simmental, Brown Swiss and crossbreed) bulls in Bolvadin district of Afyon province of Turkey, and detected a high predictive accuracy of % R^2 87.82 that the effect of BL, chest circumference, and back rump height, on FLW trait was important. The estimation of BW in indigenous beetal goat of Pakistan through head girth, neck length, diagonal body length, belly sprung, shank circumference and rump height input variables were reported by Eyduran *et al*. (2017) in the scope of CART, CHAID, RBF, MLP1, MLP2, MR modeling. We found better goodness of fit criteria (r, RMSE, AIC, MAD, SD ratio) in terms of CHAID algorithm in comparison with those recorded by Eyduran *et al*. (2017).

The difference in the results may be due the use of different sheep breeds, different body measurements, ages, environmental conditions, managerial factors, and the statistical techniques used in the study. However, it is recommended for further investigators that the predictive performances of the evaluated statistical methods should be used for different sheep breeds and studies with a large population, large number of sheep breeds and efficient factors in generalization of the results obtained from the current study.

Table II. Predictive performance of CHAID and MLR.

Methods	R^2 (%)	Adj-R^2 (%)	r	SD ratio	RMSE	RAE	MAD	CV (%)	AIC
CHAID	79.38	78.70	0.89	0.45	3.15	0.12	2.41	13.28	638.50
MLR	73.27	72.50	0.85	0.51	3.58	0.14	2.74	15.11	708.28

CONCLUSION

In conclusion, we found that for prediction of live body weight in Thalli sheep, CHAID performed better and more accurate as compared to the MLR model due to the higher Pearson correlation coefficient, R^2, Adj-R^2 and lower SD ratio, RMSE, RAE, MAD, CV%, and AIC. Although both CHAID and MLR models can remarkably predict live body weights very close to the actual values, performances of CHAID algorithm for prediction of live body weights applying body measurements of Thalli sheep was higher and more precise. Therefore, it is possible to apply CHAID algorithm, instead of traditional procedures for prediction of actual body weight using body measurements. The researchers may also use these results for comparison purposes and may be uses as a reference for next studies.

ACKNOWLEDGMENTS

We wish to acknowledge and thank the farm of sheep owners who allowed us to measure the different morphological measurements of Thalli sheep.

Conflict of interest statement

The authors have declared no conflict of interest.

REFERENCES

Akaike, H., 1973. Information theory and an extension of the maximum likelihood principle. In: *Second international symposium on information theory.* Akademiai Kiado, Hunary, pp. 267-281.

Akin, M., Hand, C., Eyduran, E. and Reed, B.M., 2017. Use of RSM and CHAID data mining algorithm for predicting mineral nutrition of hazelnut. *Pl. Cell Tissu. Organ Cult.*, **128**: 303-316. https://doi.org/10.1007/s11240-016-1110-6

Aksahan, R. and Keskin, I., 2015. Determination of the some body measurements effecting fattening final live weight of cattle by the regression tree analysis. *Selcuk J. agric. Sci.*, **2**: 53-59.

Ali, M., Eyduran, E., Tariq, M.M., Trink, C., Abbas, F., Bajwa, M.A., Baloch, M.H., Nizamani, A.H., Waheed, A., Awan, M.A., Shah, S.H., Ahmad, Z. and Jan, S., 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 J. Zool.*, **47**: 1579-1585.

Arsalan, M., Ullah, M.A. and Waheed, A., 2020. Application of CHAID algorithm for the identification of morphological traits of indigenous sheep body weight. *Proc. Pak. Acad. Sci.*, **57**: 75-80.

Behzadi, M.R.B. and Aslaminejad, A.A., 2010. A comparison of neural network and nonlinear regression predictions of sheep growth. *J. Anim. Vet. Adv.*, **9**: 2128-2131. https://doi.org/10.3923/javaa.2010.2128.2131

Berge, 1977. On the estimation of weight and increase of weight by means of the chest girth in Norwegian red cattle at the Agricultural University at As, Norway in the years 1972 and 1974. *Acta Agric. Scand.*, **27**: 65-66. https://doi.org/10.1080/00015127709435110

Buvanendra, V., Umoh, J.E. and Abubakar, B.Y., 1980. An evaluation of body size as related to weight of three West African breeds of cattle in Nigeria. *J. agric. Sci.*, (*Cambridge*), **95**: 219-224. https://doi.org/10.1017/S0021859600029476

Celik, S. and Yilmaz, O., 2017. Comparison of different data mining algorithms for prediction of body weight from several morphological measurements in dogs. *J. Anim. Pl. Sci.*, **27**: 57-64.

Celik, S., Eyduran, E., Karadas, K. and Tariq, M.M., 2017. Comparison of predictive performance of data mining algorithms in predicting body weight in Mengali rams of Pakistan. *R. Bras. Zootec.*, **46**: 863-872. https://doi.org/10.1590/s1806-92902017001100005

Chitra, R., Rajendran, S., Prasanna, D. and Kirubakaran, A., 2012. Prediction of body weight using appropriate regression model in adult female Malabari goat. *Vet. World.*, **5**: 409-411. https://doi.org/10.5455/vetworld.2012.409-411

Eyduran, E., Waheed, A., Tariq, M.M., Iqbal, F. and Ahmad, S., 2013. Prediction of live body weight from morphological characteristics of commercial goat in Pakistan using factor and principal component scores in multiple linear regression. *J. Anim. Pl. Sci.*, **23**: 1532-1540.

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 by means of body measurements in the indigenous beetal goat of Pakistan. *Pakistan J. Zool.*, **49**: 257-265. https://doi.org/10.17582/journal.pjz/2017.49.1.257.265

Goonerwardene, L.A. and Sahaayuruban, P., 1983. Analysis of body measurements and prediction of body weight in crossbred Lanka bulls. *Proc. 5th World Conf. Prod.*, **2**: 27-28.

Kass, G.V., 1980. An exploratory technique for

investigating large quantities of categorical data. *J. R. Stat. Soc. Ser. C Appl. Stat*., **29**: 119-127. https://doi.org/10.2307/2986296

Khan, B.B., Iqbal, A. and Mustafa, M.I., 2003. *Sheep and goat production*. Department Livestock Management, University of Agriculture, Faisalabad, Pakistan.

Khan, M.A., Muhammad, F., Ahmad, R., Nawaz, G., Rahimullah and Zubair, M., 2006. Relationship of body weight with linear body measurements in goats. *J. Agric. Biol. Sci*., **1**: 51-54.

Khan, M.A., Tariq, M.M., Eyduran, E., Tatliyer, A., Rafeeq, M., Abbas, F., Rashid, N., Awan, M.A. and Javed, K., 2014. Estimating body weight from several body measurements in Harnai sheep without multicollinearity problem. *J. Anim. Pl. Sci*., **24**: 120-126.

Khorshidi-Jalali, M.K., Mohammadabadi, M.R., Esmailizadeh, A., Barazandeh, A. and Babenko, O.I., 2019. Comparison of artificial neural network and regression models for prediction of body weight in raini cashmere goat. *Iran. J. Appl. Anim. Sci*., **9**: 453-461.

Kovalchuk, I.Y., Mukhitdinova, Z., Turdiyev, T., Madiyeva, G., Akin, M., Eyduran, E. and Reed, B.M., 2017. Modeling some mineral nutrient requirements for micropropagated wild apricot shoot cultures. *Pl. Cell Tissu. Organ Cult*., **129**: 325-335. https://doi.org/10.1007/s11240-017-1180-0

Mohammad, M.T., Rafeeq, M., Bajwa, M.A., Awan, M.A., Abbas, F., Waheed, A., Bukhari, F.A. and Akhtar, P., 2012. Predicting of body weight from body measurements using regression (RT) method for indigenous sheep breeds in Balochistan, Pakistan. *J. Anim. Pl. Sci*., **22**: 20-24.

Olfaz, M., Tirink, C. and Önder H., 2019. Use of CART and CHAID algorithms in Karayaka sheep breeding. *Kafkas Univ. Vet. Fak Derg*., **25**: 105-110.

Ruhil, A.P., Raja, T.V. and Gandhi, R.S., 2013. Preliminary study on prediction of body weight from morphometric measurements of goats through ANN models. *J. Ind. Soc. Agric. Stat*., **67**: 51-58.

Salehi, F., Lacroix, R. and Wade, K.M., 1998. Improving dairy yield predictions through combined record classifiers and specialized artificial neural networks. *Comput. Electron. Agric*., **20**: 199-213. https://doi.org/10.1016/S0168-1699(98)00018-0

Seifemichael, M., Kefelegn, K., Negassi, A. and Banerjee, A.K., 2014. Variability in linear body measurements and their application in predicting body weight of Afar goats in Ethiopia. *Int. J. Interdiscip. Multidiscip. Stud*., **1**: 17-25.

Sugiura, N., 1978. Further analysis of the data by Akaike's information criterion and the finite corrections. *Commun. Stat. Theory Methods*, **7**: 13-26. https://doi.org/10.1080/03610927808827599

Takma, C., Atil, H. and Aksakal, V., 2012. Çoklu doğrusal regresyon ve yapay sinir ağı modellerinin aktasyon süt verimlerine uyum yeteneklerinin karşılaştırılması. *Kafkas Univ. Vet. Fak Derg*., **18**: 941–944. https://doi.org/10.9775/kvfd.2012.6764

Tariq, M.M., Eyduran, E., Bajwa, M.A., Waheed, A., Iqbal, F. and Javed, Y., 2012. Prediction of live body weight from testicular and morphological characteristics in indigenous Mengali sheep of Pakistan: Using factor analysis scores in multiple linear regression analysis. *Int. J. Agric. Biol*., **14**: 590-594.

Willmott, C.J. and Matsuura, K., 2005. Advantages of the mean absolute error (MAE) over the root mean square error (RMSE) in assessing average model performance. *Clim. Res*., **30**: 79. https://doi.org/10.3354/cr030079

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

Yakubu, A., 2012. Application of regression tree methodology in predicting the body weight of Uda sheep. *Anim. Sci. Biotech*., **45**: 484-490.

Yapark, M., Koycegiz, F., Kutluca, M., Emsen, E. and Ockerman, W.H., 2008. Canonical correlation analysis of body measurements, growth performance and carcass trait of red Karaman lambs. *J. Anim. Vet. Adv*., **7**: 130-136.