Pakistan J. Zool., vol. 48 (4), pp. 957-960, 2016.

Prediction of Fleece Weight from Wool Characteristics of Sheep Using Regression Tree Method (Chaid Algorithm)

Ecevit Eyduran,[1] Ismail Keskin,[2] Yakup Erdal Erturk,[3] Birol Dag,[2] Adile Tatliyer,[4] Cem Tirink,[1] Rıfat Aksahan[5] and Mohammad Masood Tariq[6]*

[1] *Department of Animal Science, Agricultural Faculty, Igdir University, Igdir, Turkey*

[2]*Department of Animal Science, Agricultural Faculty, Selcuk University, Konya, Turkey*

[3]*Department of Agricultural Economy, Agricultural Faculty, Igdir University, Igdir, Turkey*

[4]*Department of Animal Science, Agricultural Faculty, Kahramanmaras Sutcu Imam University, Kahramanmaras, Turkey*

[5]*Dumlupinar Ministry of Food, Agriculture, and Livestock, Kutahya, Turkey*

[6]*Center for Advanced Studies in Vaccinology and Biotechnology, University of Balochistan, Quetta, Balochistan, Pakistan*

ABSTRACT

The goal of the current investigation was to predict the fleece weight from some physical wool characteristics of Akkaraman (47 heads) and Awassi ewes (108 heads) by using CHAID algorithm to obtain more flexible prediction through tree-based decision. All of the ewes were 2 years of age. The wool characteristics evaluated for the study were fleece weight (FW), staple length (SL), fiber length (FL), average number of crimps over a length of 5 cm (ANC) and wool fineness (WF). Results of the visual analysis from decision tree diagram demonstrated on the basis of CHAID algorithm that i) the Awassi sheep with both SL > 13 and FL $\leq$ 15 produced the heaviest FW average, ii) Akkaraman sheep was the group that had the lightest FW average (1.904 kg), iii) none of all the analyzed characteristics influenced FW trait of Akkaraman sheep, iv) FL influenced solely FW of Awassi sheep with SL > 13 (Adjusted P < 0.05). To conclude, use of CHAID algorithm for the high heritable wool characteristics of economic significance could give precious information for determining sheep with high yields.

Article Information

Received 22 April 2015
Revised 15 November 2015
Accepted 12 December 2015
Available online 1 June 2016

Authors' Contribution

BD conceived and designed the study. RA collected the data. EE and IK statistically analyzed the data. EE and MMT interpreted data, and wrote the article. YEE, AT and CT helped in interpreting the data.

Key words

Awassi breed, CHAID, Akkaraman breed, Regression tree, Sheep, Wool.

INTRODUCTION

Sheep production has played a major role in animal production, particularly in rural development. Sheep is a small ruminant that has the advantage of being the multifaceted animal species for its meat, milk, and wool. Wool is multipurpose fiber used not only to produce clothing *viz.* sweaters, dresses, coats, jackets, pants, but also to manufacture household goods such as draperies, carpets, and blankets (Tariq *et al.*, 2013).

Just as for meat and milk products of sheep, increase in wool production is based on superior offspring with desirable genetic characters.

Fiber diameter, fiber length, fiber elasticity, breaking strength and med-fibers are some of the important characteristics of wool. The interrelationship of these parameters for proper wool production should be evaluated for all the indigenous sheep breeds for the benefit of carpet industry (Kucuk *et al.*, 2000).

Tariq *et al.* (2013) reported that the heritability estimates of autumn fleece weight, spring fleece weight, and combined fleece weight were 0.19±0.03, 0.22±0.06, and 0.18±0.04, respectively in animal model and season of shearing and location of flocks had a significant impact on fleece weight (P<0.05), but birth type and sex were insignificant factors for fleece characteristics. Colakoglu and Ozbeyaz (1999) investigated greasy fleece weight and wool characteristics on Akkaraman and Malya sheep breeds at different ages. Principally, most of the studies in recent years have focused on the effect of some environmental factors (especially age) on the wool-related characteristics in some sheep breeds (Kucuk *et al.*, 2000; Yildiz and Denk, 2006; Garip *et al.*, 2010; Ozturk and Odabasioglu, 2011; Pesmen and Yardimci, 2012). Contrary to those previous reports, it is necessary that the relationships among the wool characteristics be scrutinized.

Multiple linear regression and ridge regression statistical methods to be utilized in examining the wool-related relationships can be adversely impressed by the high correlations among explanatory variables. The adverse impact is multicollinearity and causes to be misinterpretations in sheep breeding and industry.

---

* Correspondence address: tariqkianiraja@hotmail.com
0030-9923/2016/0004-0957 $ 8.00/0

However, more convenient and efficient statistical procedures are required for determining relationship between wool characteristics. The regression tree method (based on CHAID algorithm) is a tree-based non-parametric method which is eloquently interpretable due to visual form, and in which the sheep with similar characteristics will be clustered. The important advantages for the regression tree method are that no assumption is required about the distribution of explanatory variables and that the method is not influenced by the outliers (Mendes and Akkartal, 2009).

Although regression tree method has recently become popular in animal science (Eyduran *et al.,* 2008, 2013; Bakir *et al.*, 2010; Topal *et al.*, 2010; Tariq *et al.*, 2012; Yilmaz *et al.*, 2013; Khan *et al.*, 2014), it has not yet been used to determine the relationships among wool characteristics and prediction of fleece weight. The decision tree method can be considered as a valuable tool for phenotypic characterization and quality classification of the sheep with regards to their wool characteristics. The fleece weight is linked to high heritable wool characteristics, which are indirect selection criteria for wool breeding. The objectives of the current investigation were to predict the fleece weight from some wool characteristics of the sheep reared at Gözlü State Farm located in the central Anatolia region of Turkey with a regression tree method (CHAID algorithm).

MATERIALS AND METHODS

The study material consisted of wool of 47 heads of Akkaraman and 108 heads of Awassi ewes raised in the State Farm of Gözlü, Turkey. All of the ewes were with 2 years of age. Fleece weight (FW), staple length (SL), fiber length (FL), average number of crimps over a length of 5 cm (ANC) and wool fineness (WF) were investigated as physical wool properties. The data of the present study were taken from Ilhan *et al.* (2009) with the objective to introduce CHAID algorithm.

CART, CHAID, Exhaustive CHAID, MARS, Random Forest, and Machine Learning (Bayesian, Support Vectors, and Nearest Neighbor) are advanced data mining algorithms (Nisbet *et al.*, 2009). For satisfying momentous segmentations, CHAID (Chi-square Automatic Interaction Detector) tree-based segmentation algorithm with merging, splitting, and stopping stages produce multi-way splits, although CART yields binary splits in next stages (Nisbet *et al.*, 2009). Bonferroni adjustment was made for estimating adjusted P values in CHAID algorithm (Ali *et al.*, 2015).

In the present work, CHAID algorithm was employed to predict FW from other continuous variables (SL, FL, ANC, and WF) and a discrete-nominal variable (Breed) on the basis of parent node: 20 and child node: 10. CHAID algorithm utilizes the F test for significance test in the estimation of continuous response (target) variable (Nisbet *et al.*, 2009). All the statistical evaluations were performed with SPSS 22 package program (Anonymous, 2014).

RESULTS AND DISCUSSION

The solely bivariate relationships of predictors which are linked to the target variables such as milk, wool, and meat may lead to statistically false interpretations which were assumed to. Even the multiple linear regression analysis (MLRA) provides satisfactory interpretation; it can give faulty interpretation due to the biased parameter estimation. The Ridge regression, to some extent, can produce more accurate results than MLRA. When comparing the previous statistical methods, a decision tree based on the CHAID tree-based segmentation algorithm is suggested to be more efficient approach for purposeful segmentations. Due to these reasons, in the present investigation, the regression tree diagram drawn with the goal of predicting FW is visually displayed in Figure 1. The decision tree diagram exhibited openly that SL, FL, and breed (Adj P=0.000) were significant predictors in the prediction of FW, but the insignificant predictors such as WF and ANC were removed from regression tree diagram (Fig. 1). From Figure 1, it was seen markedly that the most prominent predictor on predicting FW was estimated to be breed, which was followed by SL and FL, respectively.

Root node (Node 0) that is the node at the top of the decision tree diagram was involved in all the 155 head sheep in the prediction of FW. Descriptive statistics (mean± standard deviation) for Node 0 were found to be 2.536 (S=0.749) kg. Breed had much more significant impact on FW when compared to SL and FL in the decision tree prediction model (Adjusted P = 0.000, F=69.125, df1=1 and df2=153). Node 0 was divided into two new child nodes (Nodes 1 and 2) in breed factor. In the regression tree diagram, Node 1 represented the group of only Akkaraman breed sheep (n=47), while Node 2 was confirmed to be the group of solely Awassi breed sheep (n=108). In respect of Figure 1, Akkaraman FW average (1.904, S=0.727) was lower markedly than the average of Awassi breed sheep (2.811, S=0.574), respectively. None of all the predictors affected FW characteristics of Akkaraman sheep, inferring that separation process for the subsequent stages of the tree diagram on the basis of the CHAID algorithm would be stopped in Node 1, which was defined as one of terminal nodes in the tree diagram.

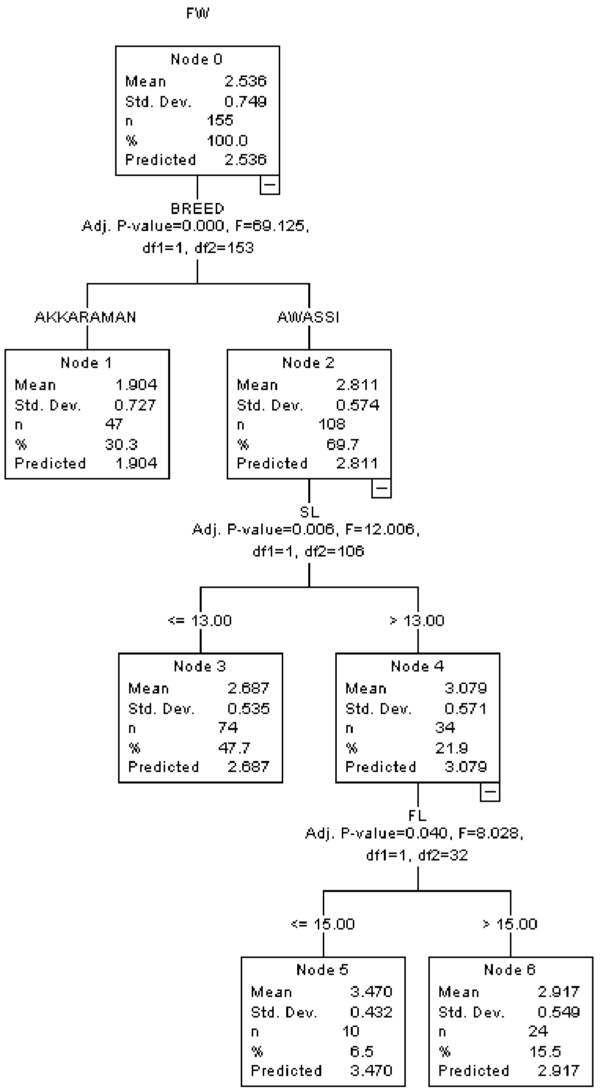

Fig. 1. Regression tree diagram for the prediction of fleece weight.

In next stage of the CHAID algorithm, Node 2 (only Awassi sheep) that was exposed to the new separation in SL predictor comprised into two new CHAID Nodes 3 and 4, respectively. As was seen nakedly from Figure 1, the SL predictor statistically affected purely FW characteristic of Awassi sheep (Adjusted P = 0.006, F=12.006, df1=1, and df2=106). The group of Awassi sheep with SL $\leq$ 13 was assigned to Node 3, another terminal node in the decision tree diagram. Node 4 was appeared from the tree diagram to be the group of Awassi sheep with SL > 13. FW averages of Nodes 3 and 4 were 2.687 (S=0.535) kg and 3.079 (S=0.571) kg, respectively. Node 4 was heavier in FW than Node 3.

In next stage, Node 4 (n=34) was partitioned into two new child nodes (Nodes 5 and 6) in terms of FL variable. Only FL influenced FW of Awassi sheep with SL > 13 (Adjusted P = 0.040, F=8.028, df1=1, and df2=32). Node 5 was perceptibly exhibited to be the group of Awassi sheep with both SL > 13 and FL $\leq$ 15. However, Node 6 exemplified the group of Awassi sheep with both SL > 13 and FL > 15. Average FW values for Nodes 5 (n=10) and 6 (n=24) were estimated to be 3.470 (S=0.432) and 2.917 (S=0.549), respectively. The FW average of the sheep in Node 5 was found, as appeared distinctly from Figure 1, to be heavier in FW as compared to the average of those in Node 6.

The visual results from decision tree diagram revealed perspicuously that Nodes 1, 3, 5 and 6 due to the stopping rules of the CHAID algorithm were terminal nodes, respectively. Additionally, the heaviest FW average was acquired from Awassi sheep with both SL > 13 and FL $\leq$ 15 in Node 5, although the lightest FW average (1.904 kg) was taken from the only group of Akkaraman sheep, unsurprisingly. It was, therefore, observed that Akkaraman sheep was lighter in FW than Awassi sheep (1.904 vs. 2.811 kg).

The present results could not compare with those reported by the earlier investigators due to sources of genetic (breed) and non-genetic (managerial conditions) variation, and more specially, the use of discrepancy statistical methods, characteristics, and factors included in all the studies.

From the point of sheep breeding, an efficient characterization of sheep breeds, depending upon the studied traits *viz.* wool, meat, and milk is associated with the selection of accurate predictors for CHAID algorithms together with the substantiation of genetic analysis in the past. Study's materials, both Akkaraman and Awassi breeds have coarse mixed fleece. In these wool types, SL and FL are major determinants of FW. Furthermore SL only depends on FL, because the fibers combine to form staples. It can be concluded that the coarse mixed fleece weight greatly determined by FL. On the contrary, fleece staples of the fine wool breeds are formed uniform fibers in respect of fiber diameter, fiber length, fiber crimp.

The present study illustrated how to interpret the results from CHAID algorithm used for the data of wool characteristics. CHAID algorithm also can be used for further studies related to milk and meat properties and may help in genetic selection. On the other hand, in the controversial cases, suspicious meats in the markets may be identified by using this method.

In order to well-understand the mechanism the investigators should conduct further studies on much larger and different sheep populations with the support of advanced molecular techniques via CHAID algorithm.

Conflict of interest statement

Authors have declared no conflict of interest.

REFERENCES

Ali, M., Eyduran, E., Tariq, M.M., Tirink, 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.

Anonymous, 2014. SPSS statistical package program (Ver: 22).

Bakir, G., Keskin, S. and Mirtagioglu, H., 2010. Determination of the effective factors for 305 days milk yield by regression tree (RT) method. *J. Anim. Vet. Adv.,* **9**: 55-59.

Colakoglu, N. and Ozbeyaz, C., 1999. Comparison of some production traits in Malya and Akkaraman sheep. *Tr. J. Vet. Anim. Sci.,* **23**: 351-360.

Eyduran, E., Karakus, K., Keskin, S. and Cengiz, F., 2008. Determination of factors influencing birth weight using regresion tre (RT) method. *J. appl. Anim. Res*., **34**: 109-112.

Eyduran, E., Yilmaz, I., Kaygisiz, A. and Aktas, Z.M., 2013. An investigation on relationship between lactation milk yield, somatic cell count and udder traits in first lactation Turkish saanen goat using different statistical techniques. *J. Anim. Pl. Sci*., **23**: 956-963.

Garip, M., Coskun, B., Polat, E.S., Yilmaz, A., Tekin, M.E., Caglayan, T. and Kilic, N., 2010. Wool properties in Kangal Akkaraman sheep. *Eurasian J. Vet. Sci*., **26**: 93-99.

Ilhan, F., Keskin, I. and D, B., 2009. *Classification of physical properties of Akkaraman and Awassi wools using discriminant analysis*. VI. Animal Sci. Congress. 24-26 June, Erzurum, Turkiye.

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.

Kucuk, M., Yilmaz, O. and Ates, C.T., 2000. The evaluation of Morkaraman, Hamdani, Karagul wool for carpet wool type. *Van Vet. J.,* **11**: 54-59.

Mendes, M. and Akkartal, E., 2009. Regresion tree analysis for predicting slaughter weight in broilers. *Italian J. Anim. Sci.,* **8**: 615-624.

Nisbet, R., Elder, J. and Miner, G., 2009. *Handbook of statistical analysis and data mining applications*. Canada.

Ozturk, Y. and Odabasioglu, F., 2011. The investigation of yield and morphological characteristics of Hamdani sheep in Van area; I. various yield characteristics of sheep. *Van Vet. J.,* **22**: 75-80.

Pesmen, G. and Yardimci, M., 2012. Fleece qualities of Menemen sheep: I. Morphologic and physical qualities. *Eurasian J. Vet. Sci.,* **28**: 99-105.

Tariq, M.M., Rafeq, M., Bajwa, M. A., Awan, M. A., Abbas, F., Waheed, A., Bukhari, F.A. and Akhtar, P., 2012. Prediction of body weight from body measurements using regresion tree (RT) method for indigenous sheep breeds in Balochistan, Pakistan. *J. Anim. Pl. Sci.,* **2**: 20-24.

Tariq, M.M., Bajwa, M.A., Javed, K., Rafeeq, M., Awan, M.A., Rashid, N., Waheed, A., Amin, S., Khan, M.A. and Bukhari, F.A., 2013. Assesment of wool characteristics of Mengali sheep of Balochistan. *J. Anim. Pl. Sci.,* **23**: 721-726.

Topal, M., Aksakal, V., Bayram, B. and Yaganoglu, M., 2010. An analysis of the factor affecting birth weight and actual milk yield in Swedish red cattle using regression tree analysis. *J. Anim. Pl. Sci.,* **20**: 63-69.

Yildiz, N. and Denk, H., 2006. Some of production traits of Akkaraman ewes raised by farmers in Van region I. reproductive and milk production traits. *Firat Univ. Vet. J. Hlth. Sci.,* **20**: 21-27.

Yilmaz, I., Eyduran, E. and Kaygisiz, A., 2013. Determination of non-genetic factors influencing birth weight using regression tree method in Brown-Swiss cattle. *Canadian J. appl. Sci.,* **1**:382-387.