



Multivariate modeling to estimate the composition of carcass tissues of Santa Inês sheep

Marilene dos Santos Maciel^{1*}, Janaína Kelli Gomes Arandas¹, Francisco Fernando Ramos de Carvalho¹, George Rodrigo Beltrão da Cruz², Roberto Germano Costa³, Neila Lidiany Ribeiro⁴ and Maria Norma Ribeiro¹

¹Departamento de Zootecnia, Universidade Federal Rural de Pernambuco, Rua Dom Manuel de Medeiros, s/n, Dois Irmãos, 51171-900, Recife, Pernambuco, Brazil. ²Departamento de Ciência Animal, Centro de Ciências Humanas, Sociais e Agrárias, Universidade Federal da Paraíba, Campus Universitário III, Bananeiras, Paraíba, Brazil. ³Departamento de Zootecnia, Universidade Federal da Paraíba, Areia, Paraíba, Brazil. ⁴Universidade Federal de Campina Grande, Campina Grande, Paraíba, Brazil. *Author for correspondence. E-mail: marilenemaciel123@hotmail.com

ABSTRACT. The purpose of this study was to establish a multivariate model using two complementary multivariate statistical techniques: Factor Analysis and Stepwise Multiple Regression, to predict tissue composition through carcass characteristics of Santa Inês sheep. The data was obtained from 82 Santa Inês sheep under confinement. The predictor variables were carcass characteristics related to weight, yield, morphometric measures and meat cuts. The use of latent variables from factor analysis in multiple regression models eliminates the problem of multicollinearity of the explanatory variables, improving the accuracy of interpretation of results by proposing a better fit of the mathematical model. However, the coefficient of determination (R^2) values were moderate for muscle proportion and total fat, and low for bone proportion, indicating that more appropriate independent variables should be used to better predict the proportion of tissues in Santa Inês sheep.

Keywords: multivariate analysis; carcass dissection; factor scores; morphometric measures.

Received on July 30 2022.
Accepted on January 30 2023.

Introduction

Carcass evaluation is a key process in determining the value and quality characteristics of production animals destined for slaughter. To a large extent, commercial value is related to carcass yield and quality. As described by Ekiz, Baygul, Yalcintan, and Ozcan (2020), carcass yield and composition (proportions of muscle, fat, and bone) are important determinants of carcass quality due to the high variability observed in these characteristics and their obvious effects on commercial value. However, to determine composition more accurately, a total or partial dissection of these components is required, which is an expensive and time-consuming method.

Therefore, multiple regression analysis is a commonly used prediction model for interpretation between a dependent variable and two or more independent variables, however, this method has some disadvantages. The development of multiple regression models using independent variables with high correlations may present limitations in their inference and accuracy, and are likely to have serious effects on the estimates of regression coefficients and the overall applicability of the estimated model (Gomes et al., 2013), due to the problem of multicollinearity.

To avoid this problem, studies have been carried out using orthogonal factor scores (latent variables) present in multivariate factor analysis. Çelik et al. (2018) used factor scores to evaluate the influence of carcass parts weights on total weight in turkeys. Daskiran, Keskin, and Bingol (2017) determined the relationship between daily milk production and udder characteristics in goats through factor scores. Önk, Sari, and Gürcan (2018), Tahtali (2019) and Tariq et al. (2012) also used factor scores to estimate body weights in lambs. In a recent study, factor analysis and multiple stepwise regression were used to predict carcass characteristics, carcass cuts, internal fat, viscera and loin eye area from body measurements of crossbred Boer goats (Macena et al., 2022).

The prediction of body and carcass composition of ruminants was first proposed by Hankis and Howe (1946), who showed that the chemical composition of the section of the 9, 10 and 11th ribs was significantly

correlated with carcass composition in beef cattle. This triggered subsequent studies that showed that different cuts and carcass measures can be effective predictors in the evaluation of body composition and overall carcass of ruminants (Fernandes et al., 2008; Lambe et al., 2009; Marcondes, Tedeschi, Valadares Filho, & Chizzotti, 2012; Ribeiro & Tedeschi, 2012).

Therefore, it was hypothesized that orthogonal factor scores arising from the combination of different carcass traits can produce reliable predictions of the tissue composition of Santa Inês sheep. Based on the information above, the purpose of this study was to establish a multivariate model using two complementary multivariate statistical techniques: Factor Analysis and Stepwise Multiple Regression, to predict the tissue composition in Santa Inês sheep, using the characteristics of weight, yield, morphometric measures and meat cuts as independent variables.

Material and methods

Experiment and animals

The experiments were carried out in the Goat and Sheep breeding sector of the Human, Social and Agriculture Sciences Center of the Federal University of Paraíba, which is located in the city of Bananeiras, state of Paraíba, Brazil. A total of 82 Santa Inês sheep were used from two experiments that were carried out to determine carcass characteristics and meat quality under confinement. The research protocols of the two experiments were approved by the Ethics Committee of the Federal University of Paraíba.

The experiment 1 aimed to evaluate different levels of cactus pear inclusion (*Opuntia ficus-indica*, Mill) in the diet and restriction of voluntary water intake on performance, carcass characteristics and meat quality of Santa Inês sheep. The experiment 2 aimed to evaluate carcass characteristics and meat quality in Santa Inês sheep fed with increasing levels of guava agro-industrial waste (*Psidium guajava* L.) in the diet. The main information of the experiments is shown in Table 1.

Table 1. Main information of the experiments.

Characteristics	Experiment 1	Experiment 2
Number of animals	42.0	40.0
Initial age, days	180.0	120.0
Initial weight, kg	21.6 ± 2.2	21.3 ± 2.2
Slaughter weight, kg	28.9 ± 2.6	34.2 ± 3.4
Cold carcass weight, kg	14.1 ± 1.3	15.6 ± 1.8

Slaughter procedures and carcass characteristics

Slaughter was performed according to the current RIISPOA (Brasil, 2000) norms; the animals were stunned by captive dart pistol, with stunning followed by bleeding for four minutes, through carotid and jugular sections. The blood was collected in a previously weighed container for later weighing.

After skinning and evisceration, the head (section at the atlanto-occipital joint) and legs (section at the metacarpal and metatarsal joints) were removed and the hot carcass weight (HCW) was recorded. The internal components of the pelvic, abdominal and thoracic cavities were extracted and their weights were recorded. After obtaining the hot carcass weight (HCW), the carcasses were taken to the cold chamber, at an average temperature of 4°C, in which they remained for 24 hours suspended on hooks by the tendon of the gastrocnemius muscle, and then the cold carcass weight (CCW) was obtained, according to the methodology of Cezar and Souza (Cezar & Souza, 2007).

The gastrointestinal tract (GIT) was weighed full and empty for determination of empty body weight (EBW, kg) and biological yield (BY%) = $HCW/EBW \times 100$. Kidneys and perirenal fat were removed and subtracted from HCW and CCW to calculate hot carcass yield (HCY, %) = $HCW/LWS \times 100$, cold carcass yield (CCY, %) = $CCW/LWS \times 100$ and cooling loss (CL, %) = $(HCW - CCW)/HCW \times 100$, according to Cezar and Souza (Cezar & Souza, 2007).

After the cooling period, the carcasses were sectioned in half and the half-carcasses were weighed. In the left half-carcass the internal and external length, leg length, thorax perimeter, croup perimeter, thorax depth, thorax width and croup width were measured, according to the methodology proposed by Cezar and Souza (2007). The carcass compactness index (CCI) was also calculated through the equation $CCI (kg/cm) = CCW/\text{carcass internal length}$, according to methodology proposed by Cezar and Souza (Cezar & Souza, 2007).

After the carcasses were longitudinally divided, the half carcasses were sectioned into five anatomical regions that composed the commercial cuts, according to the methodology adapted from Cezar and Souza (2007), divided into neck, shoulder, rib, loin and leg. Then, the individual weight of each cut from the left half carcass was recorded to calculate their proportion in relation to the sum of the reconstituted half carcass, thus obtaining the yield of the carcass cuts.

For further evaluation of tissue composition, the left leg from each animal was packed in high-density polyethylene bags and frozen at -18°C . In order to determine the tissue composition, these pieces were then dissected, according to the methodology described by Brown and Williams (1979), being previously gradually thawed and kept at a temperature of approximately 4°C for 24 hours.

With the use of scalpel, tweezers and scissors, the following tissue groups were separated: subcutaneous fat, intermuscular fat (all fat located below the deep fascia, associated with muscles), muscle (total weight of muscles dissected after complete removal of all adhered intermuscular fat), bone (total weight of the leg bones), and other tissues (all unidentified tissues, composed of tendons, glands, nerves, and blood vessels). The weights and yields of the dissected tissues were obtained by dissecting the leg, and the percentage of tissue components was calculated in relation to the reconstituted weight of the leg after dissection.

Statistical analyses

Descriptive statistics (mean, standard deviation, variance, minimum and maximum values) were determined for all variables. Pearson's analysis was used to determine the correlation coefficient of dissected tissue compositions (muscle, bone, and fat) with the independent variables. Regressions were carried out through PROC REG of SAS® OnDemand for Academics.

The effectiveness of the multiple regression analysis was determined through coefficient of determination (R^2), mean square error (MSE), variance inflation factor (VIF) as indicator of multicollinearity, and Mallows' C_p statistic. The VIF is an indicator of multicollinearity and indicates how much a regression coefficient is increased due to correlations between predictors in the model.

Mallows' C_p (Mallows, 2000) is a measure of quality of fit, which is often used to evaluate the regression model (Miyashiro & Takano, 2015). Mallows' C_p is given by (Equation 1):

$$\frac{RSS}{\sigma^2} + 2p - n \quad (1)$$

where RSS is the residual sum of squares, σ^2 is the residual variance, p is the number of parameters in the model (including the intercept) and n is the number of variables. The goal is to find the best model involving a subset of predictors. Thus, only those models that have C_p values close to the number of parameters (including the intercept) should be considered as a desired criterion for selecting a subset of predictors (Mallows, 2000).

Multiple regression analysis was used to estimate the dissected tissue compositions from different carcass characteristics. However, in the set of independent variables there may be variables that have little influence on the dependent variables; thus, the stepwise procedure was used to select which variables have the most influence on the dependent variables and thus can decrease the number of variables to compose the model equation (Alves, Lotufo, & Lopes, 2013). According to Senra, Nanci, Mello, and Meza (2007), this procedure is based on the observation that some variables have little contribution to the average efficiency of the model; therefore, once identified, they can be removed from the model.

The stepwise multiple regression analysis was performed using the model (Equation 2):

$$Y = \alpha + \beta_1 X_1 + \beta_2 X_2 + \dots + \beta_n X_n + e \quad (2)$$

where Y is the dependent or response variable; α is the intercept of the regression equation, β_1 , β_2 and β_n are regression coefficients of variables X_1 , X_2 and X_n that are the independent or explanatory variables and e is the residual random error. The criterion used for entry and permanence of an independent variable in the model was $p > 0.05$.

In multiple regression analysis, an estimation method based on the factor scores from factor analysis can be used to eliminate the limitations caused by the multicollinearity problem among the independent variables. The main purpose of factor analysis is to allow understanding and interpreting the supposed relationship between multiple variables and to represent these multiple variables in a minimum number of factors (latent variables) required for the maximum variance represented by the original variables.

Bartlett's test and the Kaiser-Meyer-Olkin (KMO) test are applied to test the divisibility of the correlation matrix into factors (Tahtali, 2019). If the null hypothesis is rejected according to the results of Bartlett's test, the factor analysis can be continued. An index below 0.5 with the KMO test indicates that the relationship between pairs of variables cannot be explained by other variables (Çelik et al., 2018), showing inadequacy. Orthogonal Varimax rotation was employed to improve the interpretation of the extracted factors.

The PROC FACTOR procedure of the statistical software SAS® OnDemand for Academics was used for the retention of the factors through principal component analysis and the choice of the number of factors through Kaiser's (1974) criterion, which considers eigenvalues ≥ 1 as significant.

The model used in the analysis was (Equation 3):

$$\begin{aligned} X_1 &= a_{11}F_1 + a_{12}F_2 + \dots + a_{1m}F_m + \varepsilon_1 \\ X_2 &= a_{21}F_1 + a_{22}F_2 + \dots + a_{2m}F_m + \varepsilon_2 \\ X_p &= a_{p1}F_1 + a_{p2}F_2 + \dots + a_{pm}F_m + \varepsilon_p \end{aligned} \quad (3)$$

The factor model was built by the factors $F_1, \dots, F_m, m \leq p$. Where: X_1, X_2 and X_p = studied variables; a = factor loadings; $F = F_1, F_2, \dots, F_m$ are the non-correlated factors; $\varepsilon = \varepsilon_1, \varepsilon_2, \dots, \varepsilon_p$ are variables with mean 0 and variance 1.

Thus, the multiple regression analysis was also used to estimate carcass tissue composition from the extracted factors, according to the model: $Y = \alpha + \beta_1F_1 + \beta_2F_2 + \dots + \beta_nF_n + e$ (Equation 4), in which Y is the dependent/response variable; α is the intercept of the regression analysis; β_1, β_2 and β_n are the regression coefficients of scores F_1, F_2 and F_n that are the explanatory variables or factors, and e = residual random error.

Results and discussion

Descriptive statistics of dependent and independent variables

The proportions of tissues dissected from the sheep carcasses are shown in Table 2. The proportions of muscle, bone and total fat were 68.14, 18.87 and 9.36, respectively. These values were consistent with those reported previously for carcasses of Santa Inês sheep (Cardoso et al., 2021; Fernandes et al., 2021). Similar tissue proportions were also reported for sheep with no defined racial pattern by Lima Júnior et al. (2017).

Table 2. Descriptive statistics of dependent variables.

Variables	$\mu \pm sd$	Variance	Min.	Max.
Muscle proportion (%)	68.14 \pm 3.88	15.04	58.87	77.64
Bone proportion (%)	18.87 \pm 2.26	5.09	15.28	27.81
Total fat proportion (%)	9.36 \pm 2.85	8.14	4.10	17.03

$\mu \pm sd$ = mean \pm standard deviation.

Descriptive statistics of the independent variables used to predict carcass tissue composition are shown in Table 3. The presented variables are related to weight, yield and meat cut characteristics. Studies have shown that different carcass parts and measures can be accurate predictors of body and carcass composition in ruminant animals, considering them as independent variables (Lambe et al., 2009; Marcondes et al., 2012; Ribeiro & Tedeschi, 2012).

Correlation coefficients between carcass characteristics and tissue composition

The correlation coefficients between the proportions of carcass tissues and carcass characteristics are shown in Table 4.

Muscle proportion had positive correlation coefficients with traits related to yield variables (BY, HCY, CCY, RIBY, NECY and SHOY), LP and CCI. The strongest correlations of muscle proportion were obtained with the measure variables CIL ($r = -0.590$), LL ($r = -0.588$), CW ($r = -0.688$), TW ($r = -0.661$) and with the meat cuts LOI ($r = -0.570$) and SHO ($r = -0.672$). Ekiz et al. (2020) evaluated Gokceada breed goats and found no significant result ($p > 0.05$) between the variable CIL and muscle proportion. On the other hand, the correlations of muscle proportion with certain carcass characteristics, such as CCW, CP, TP, RIB, and SHO, were not significant in this study ($p > 0.05$).

Table 3. Descriptive statistics of independent variables.

Variables	Abbreviations	$\mu \pm \text{sd}$	Variance	Min.	Max.
<i>Weight (kg)</i>					
Live weight at slaughter	LWS	31.58 \pm 4.00	15.97	22.70	42.00
Empty body weight	EBW	25.73 \pm 3.33	11.06	18.27	34.70
Hot carcass weight	HCW	15.17 \pm 1.80	3.25	11.20	20.20
Cold carcass weight	CCW	14.87 \pm 1.73	2.98	11.09	20.00
<i>Yield (%)</i>					
Biological yield	BY	59.15 \pm 3.68	13.51	48.12	73.88
Hot carcass yield	HCY	48.18 \pm 3.07	9.45	38.64	60.26
Cold carcass yield	CCY	47.27 \pm 3.29	10.82	37.08	59.60
Leg yield	LEGY	30.08 \pm 3.75	14.05	19.37	41.27
Loin yield	LOIY	12.93 \pm 2.17	4.70	7.63	20.12
Rib yield	RIBY	29.99 \pm 3.89	15.16	19.74	38.53
Neck yield	NECY	6.92 \pm 1.19	1.41	4.18	9.98
Shoulder yield	SHOY	18.81 \pm 2.44	5.93	10.90	23.78
<i>Morphometric measures (cm)</i>					
Carcass external length	CEL	56.39 \pm 3.03	9.15	48.00	65.00
Carcass internal length	CIL	60.66 \pm 4.59	21.04	51.00	71.00
Leg length	LL	34.57 \pm 2.85	8.12	28.00	39.00
Croup width	CW	17.99 \pm 3.47	12.07	12.20	26.00
Thorax width	TW	17.21 \pm 4.08	16.61	11.20	24.50
Croup perimeter	CP	57.09 \pm 5.47	29.93	49.00	91.00
Leg perimeter	LP	37.31 \pm 3.09	9.53	28.00	43.00
Thorax perimeter	TP	67.96 \pm 3.77	14.22	58.00	85.50
Carcass compactness index	CCI	0.25 \pm 0.02	0.00	0.20	0.29
<i>Cuts (kg)</i>					
Leg	LEG	2.22 \pm 0.30	0.09	1.41	3.02
Loin	LOI	0.96 \pm 0.18	0.03	0.64	1.46
Rib	RIB	2.22 \pm 0.30	0.09	1.51	3.14
Neck	NEC	0.76 \pm 0.26	0.07	0.38	1.27
Shoulder	SHO	1.39 \pm 0.17	0.03	0.99	1.83

Table 4. Pearson's correlation (r) between tissue composition and carcass variables.

Variables	Abbreviations	Muscle (%)	Bone (%)	Fat (%)
<i>Weight (kg)</i>				
Live weight at slaughter	LWS	-0.412*	-0.362*	0.501*
Empty body weight	EBW	-0.428*	-0.364*	0.525*
Hot carcass weight	HCW	-0.255*	-0.361*	0.412*
Cold carcass weight	CCW	-0.215 ^{NS}	-0.327*	0.367*
<i>Yield (%)</i>				
Biological yield	BY	0.418*	0.069 ^{NS}	-0.331*
Hot carcass yield	HCY	0.356*	0.048 ^{NS}	-0.246*
Cold carcass yield	CCY	0.407*	0.113 ^{NS}	-0.319*
Leg yield	LEGY	-0.096 ^{NS}	0.094 ^{NS}	0.033 ^{NS}
Loin yield	LOIY	-0.525*	-0.142 ^{NS}	0.494*
Rib yield	RIBY	0.222*	0.155 ^{NS}	-0.206 ^{NS}
Neck yield	NECY	0.371*	-0.091 ^{NS}	-0.235 ^{NS}
Shoulder yield	SHOY	0.307*	0.093 ^{NS}	-0.282*
<i>Morphometric measures (cm)</i>				
Carcass external length	CEL	-0.411*	-0.002 ^{NS}	0.353*
Carcass internal length	CIL	-0.590*	-0.186 ^{NS}	0.567*
Leg length	LL	-0.588*	-0.079 ^{NS}	0.486*
Croup width	CW	-0.688*	-0.313*	0.618*
Thorax width	TW	-0.661*	-0.348*	0.605*
Croup perimeter	CP	-0.204 ^{NS}	-0.199 ^{NS}	0.205 ^{NS}
Leg perimeter	LP	0.490*	0.039 ^{NS}	-0.435*
Thorax perimeter	TP	-0.180 ^{NS}	-0.271*	0.270*
Carcass compactness index	CCI	0.258*	-0.273*	-0.039 ^{NS}
<i>Cuts (kg)</i>				
Leg	LEG	-0.239*	-0.201 ^{NS}	0.324*
Loin	LOI	-0.570*	-0.324*	0.643*
Rib	RIB	0.061 ^{NS}	-0.144 ^{NS}	0.097 ^{NS}
Neck	NEC	-0.672*	-0.446*	0.746*
Shoulder	SHO	0.161 ^{NS}	-0.212 ^{NS}	0.023 ^{NS}

^{NS}p > 0.05; *p < 0.05.

Bone proportion showed significant negative correlation coefficients ($p < 0.05$) with morphometric measure and cut traits (CP, TW, TP, CCI, LOI and NEC); the other variables showed no significant correlations ($p > 0.05$) (Table 4). These results agree with Díaz et al. (2004) and Ekiz et al. (2020), who reported that bone proportion had negative and significant correlations with traits associated with weight, carcass measurements and indexes.

Except for CP, CCI, RIB, SHO, LEGY, RIBY and NECY, the other variables showed significant correlations with fat proportion ($p < 0.05$) (Table 4). Among these traits, the ones that showed the highest correlations were CW, TW, LOI and NEC, as well as in a study of Santos, Silvestre, Azevedo, and Silva (2017), who reported significant and positive correlations of fat proportion with cold carcass weight, carcass measures and meat cuts of the carcass of lactating goat kids.

Regardless of whether the correlations were positive or negative, it can be observed that most of them were significant ($p < 0.05$), indicating that these variables can be used as indicators for tissue proportion.

Prediction of carcass tissue composition using stepwise multiple regression analysis of the original variables

The regression equations for predicting tissue composition by stepwise multiple regression analysis of the original variables are presented in Table 5. According to the analysis, the independent variables CW, CCI, LOI and SHO were the best predictors of muscle proportion, explaining 63.6% of the variation in muscle proportion based on R^2 .

Table 5. Prediction equations for tissue composition according to stepwise multiple regression analysis of the original variables.

Dependent variable	Number	Equation	C_p	R^2	MSE	P Value
Muscle proportion (%)	1	$Y = 91.961 - 0.768 \times CW$	27.290	0.474	2.831	<.0001
	2	$Y = 68.314 - 0.798 \times CW + 57.885 \times CCI$	9.595	0.574	2.563	<.0001
	3	$Y = 69.817 - 0.653 \times CW + 57.293 \times CCI - 4.227 \times LOI$	7.392	0.596	2.514	<.0001
	4	$Y = 68.151 - 0.493 \times CW + 39.931 \times CCI - 9.014 \times LOI + 6.582 \times SHO$	1.438	0.636	2.401	<.0001
Bone proportion (%)	1	$Y = 21.853 - 3.940 \times NEC$	16.141	0.199	2.032	<.0001
	2	$Y = 29.708 - 31.432 \times CCI - 4.076 \times NEC$	8.044	0.287	1.930	<.0001
	3	$Y = 21.476 + 0.264 \times LL - 29.227 \times CCI - 5.984 \times NEC$	2.676	0.351	1.853	<.0001
	4	$Y = 25.638 + 0.331 \times LL - 0.117 \times TP - 23.331 \times CCI - 5.927 \times NEC$	1.411	0.379	1.824	<.0001
	5	$Y = 25.682 + 0.374 \times LL + 0.087 \times CP - 0.189 \times TP - 28.207 \times CCI - 6.420 \times NEC$	0.639	0.403	1.801	<.0001
	6	$Y = 21.476 + 0.110 \times CCY + 0.404 \times LL + 0.093 \times CP - 0.208 \times TP - 33.589 \times CCI - 6.094 \times NEC$	0.512	0.422	1.785	<.0001
Total fat proportion (%)	1	$Y = 3.052 + 8.331 \times NEC$	2.420	0.556	3.660	<.0001
	2	$Y = 5.718 + 8.137 \times NEC - 0.364 \times NECY$	0.429	0.579	3.518	<.0001

Mallows' C_p (C_p); Coefficient of determination (R^2); Mean square error (MSE); Croup width (CW); Carcass compactness index (CCI); Loin (LOI); Shoulder (SHO); Neck (NEC); Leg length (LL); Thorax perimeter (TP); Croup perimeter (CP); Cold carcass yield (CCY); Neck yield (NECY).

However, the results based on C_p values indicate that the proposed models showed a lack of fit. For acceptable models within a subset of variables, the C_p values need to be close to the number of predictors plus the constant, which indicates that the model is relatively unbiased in estimating the true regression coefficients and predicting future responses (Kazemi, Mohamed, Shareef, & Zayandehroodi, 2013). Such results may be directly related to the presence of multicollinearity among the independent variables. The variables used in the models were moderately correlated ($1 < VIF < 5$); thus, the use of carcass variables should be applied with caution, as multicollinearity is associated with unstable estimates of the regression coefficients.

In studies with these same perspectives, Díaz et al. (2004) determined that the best prediction equation for carcass muscle proportion of lactating Manchego lambs included proportion of kidney knob channel fat, fat thickness, CIL and fore cannon bone weight as independent variables. In lambs of the Churra Tensina breed, the regression equation included carcass width and kidney knob channel fat weight as independent variables to better predict muscle proportion (Carrasco, Ripoll, Panea, Álvarez-Rodríguez, & Joy, 2009). Ekiz et al. (2020) reported kidney knob and channel fat percentage as the best predictor of muscle proportion. The equations reported by Díaz et al. (2004) and Carrasco et al. (2009) obtained similar accuracy when compared to the present study (R^2 values = 0.63 and 0.58, respectively), and higher accuracy when compared to the study of Ekiz et al. (2020), who evaluated goats of the Gokceada breed, and found R^2 value = 0.21. According to Ekiz et al. (2020), the differences between the studies in terms of accuracy may be due to species differences, as well as differences in the number of independent variables allocated in the model.

The variables LL, CP, TP, CCI, and NEC were included in the equation for predicting bone proportion (Table 5). These five variables explained only 40.3% of the variation in bone proportion, which indicates that the prediction of bone proportion using this model was low. It was expected that, among the proposed models, this one would be the best for predicting bone proportion, however, based on the C_p values, the best fitted model was the one that considered only the variables LL, CCI and NEC as predictors. Although there are differences in the literature to decide which individual variable is more appropriate to be used in the prediction of animal carcass, the accuracy of the prediction in relation to R^2 has been improved especially when more than one variable is considered in the model. In the present study, it was demonstrated that adding more predictor variables and improving the R^2 does not always produce an increase in the accuracy of the obtained estimates.

In a previous study carried out to predict the bone proportion of Churra Tensina breed lambs, Carrasco et al. (2009) found out that the best obtained equation included kidney knob channel fat weight and conformation score ($R^2 = 0.51$). The equation reported by Díaz et al. (2004) included fat score and omental fat proportion to predict bone proportion of lactating lambs of the Manchego breed ($R^2 = 0.76$). On the other hand, Ekiz et al. (2020) included hind limb compactness and tail weight in the equation for predicting bone proportion ($R^2 = 0.62$).

According to the results of stepwise multiple regression analysis of the original variables, two variables (NEC and NECY) were determined to predict the proportion of total fat. This prediction equation explained 57.9% of the proportion of total fat. Díaz et al. (2004), Carrasco et al. (2009) e Ekiz et al. (2020) predicted the proportion of total fat with higher accuracies ($R^2 = 0.84, 0.73$ and 0.68 , respectively). In the study carried out by Díaz et al. (2004), the equation of the obtained model included fat variables (fat thickness, fat score and kidney knob channel fat proportion) for the prediction of fat proportion. In the study carried out by Carrasco et al. (2009), the equation model included kidney knob channel fat weight, carcass width, and carcass internal length as independent variables.

In the present study, the predicted bone proportion was less accurate (lower R^2) than the muscle and fat proportions (Table 5). On the contrary, Ekiz et al. (2020) found that the amount of explained variation was lower in the prediction equation for muscle proportion than that for bone and fat. As for the regression equations for predicting tissue compositions, it is worth mentioning that the morphometric measures were not allowed in the models except for CW, LL and TP. In a study carried out with goats of the Gokceada breed, Ekiz et al. (2020) also observed the lack of relationship between carcass size measures and tissue composition. According to Cadavez (2009), carcass measures reflect skeletal size rather than carcass tissue composition.

Based on the observed correlations between tissue composition and carcass characteristics (Table 4), it would be expected for other carcass measures to be included in the prediction equations. In that case, it is suggested the problem of multicollinearity, since moderate multicollinearity problems ($1 < VIF < 5$) were observed in the present study.

According to the present results, the prediction equations obtained by regression of the original variables seem to be insufficient to accurately predict tissue composition, since the amount of explained variation was low for bone proportion and moderate for muscle and fat proportion. On the other hand, the linear and circular carcass measures considered in the study were excluded from the models during the stepwise regression selection procedures, which indicate that these variables were poor predictors for the proportions of muscle, bone, and total fat in Santa Inês sheep.

Prediction of carcass tissue composition using stepwise regression of latent variables

Table 6 presents the results of sphericity test of Barlett and Kaiser-Meyer-Olkin (KMO), which are essential prerequisites for factor analysis. In other words, significant results from both tests show that the data are suitable for factor analysis. Considering the values obtained by Bartlett's test ($p < 0.05$) and KMO (0.702) to test the divisibility of the correlation matrix of the factors, the data were considered suitable for factor analysis.

Table 6. Sampling adequacy by Barlett's test of sphericity and the KMO (Kaiser-Meyer-Olkin) test.

	DF	325
Bartlett's sphericity test	Chi-Square	4761.91
	Pr > Chi Sq	<.0001
KMO test (Kaiser-Meyer-Olkin)	---	0.702

The values of the factor loadings obtained by Varimax rotation, communalities, eigenvalues and explained variance are described in Table 7. It is observed that retaining factors through principal component analysis

and choosing the number of factors through Kaiser's criterion made it possible to extract six factors (eigenvalues>1). These six factors were able to explain about 85.40% of the original variance of the variables. This value proved to be quite significant, and it represents little loss of information, indicating that the analysis summarized most of the information in a minimum number of factors needed to explain the maximum variance represented by the original variables.

The communality presented in Table 7 is the proportion of each variable's variability that is explained by the factors. The closer to 1, the better the variable explains the variation in the factors. Hair Jr., Black, Babin, Anderson, and Tatham (2009) describe that at least half of the variance of each variable should be taken into account, thus, using this guideline, all variables with communality lower than 0.50 present insufficient explanation. In Table 6, all variables presented communality greater than 0.50, indicating that the proportion of variability is adequate to explain the factors.

Table 7. Results of factor analysis applied to the independent variables.

Variables	Factor 1	Factor 2	Factor 3	Factor 4	Factor 5	Factor 6	C
LWS	0.762	0.122	-0.404	0.349	0.315	-0.035	0.980
EBW	0.773	0.131	-0.353	0.342	0.348	-0.032	0.978
HCW	0.783	0.078	0.074	0.466	0.368	-0.017	0.978
CCW	0.779	0.011	0.112	0.469	0.376	-0.052	0.984
BY	-0.102	-0.135	0.904	0.153	-0.033	0.033	0.871
HCY	-0.046	-0.112	0.956	0.159	0.056	0.035	0.959
CCY	-0.085	-0.215	0.939	0.130	0.045	-0.022	0.955
CEL	0.712	0.319	-0.051	0.002	0.089	-0.185	0.653
CIL	0.600	0.550	-0.184	0.080	0.361	-0.049	0.835
LL	0.617	0.598	-0.171	-0.010	0.087	-0.051	0.778
CW	0.523	0.473	-0.355	0.121	0.450	0.074	0.846
TW	0.454	0.409	-0.396	0.094	0.526	0.092	0.824
CP	0.179	-0.071	0.015	0.167	0.859	0.063	0.807
LP	-0.123	-0.774	0.091	0.116	0.171	-0.061	0.670
TP	0.329	0.137	0.162	-0.042	0.699	-0.133	0.661
CCI	0.419	-0.546	0.202	0.573	0.230	-0.047	0.898
LEG	0.302	0.220	0.113	0.431	0.156	0.770	0.956
LOI	0.235	0.697	-0.034	0.462	0.248	0.340	0.932
RIB	0.076	0.212	0.389	0.760	-0.062	0.069	0.789
NEC	0.309	0.706	-0.268	0.244	0.347	0.089	0.855
SHO	-0.163	0.028	0.109	0.863	0.131	0.178	0.833
LEGY	-0.357	0.229	-0.005	0.000	-0.153	0.872	0.963
LOIY	-0.251	0.789	-0.147	0.188	0.040	0.413	0.916
RIBY	-0.570	0.210	0.280	0.342	-0.387	0.091	0.723
NECY	-0.802	-0.015	-0.005	0.199	0.012	-0.194	0.720
SHOY	-0.791	0.016	-0.010	0.382	-0.186	0.182	0.840
Eigenvalues	9.802	4.437	4.105	1.702	1.119	1.039	---
Variance (%)	37.70	17.06	15.79	6.54	4.30	4.00	85.40

Communality (C); Live weight at slaughter (LWS); Empty body weight (EBW); Hot carcass weight (HCW); Cold carcass weight (CCW); Biological yield (BY); Hot carcass yield (HCY); Cold carcass yield (CCY); Carcass external length (CEL); Carcass internal length (CIL); Leg length (LL); Croup width (CW); Thorax width (TW); Croup perimeter (CP); Leg perimeter (LP); Thorax perimeter (TP); Carcass Compactness Index (CCI); Leg (LEG); Loin (LOI); Rib (RIB); Neck (NEC); Shoulder (SHO); Leg yield (LEGY); Loin yield (LOIY); Rib yield (RIBY); Neck yield (NECY); Shoulder yield (SHOY).

Table 8 illustrates the results of the stepwise multiple regression analysis with the new independent variables (latent) from the factor analysis.

According to the analysis, all latent variables selected through principal component analysis (Factor1, 2, 3, 4, 5 and 6) were found to be best predictors of muscle proportion, explaining 62.5% of the variation in muscle proportion. However, the variation in muscle proportion explained by these latent variables is lower than the result found in the regression analysis of the original variables (Table 5), with the best model (CW, CCI, LOI and SHO) explaining 63.7% of the variation for muscle proportion. Nevertheless, the latent variables should be considered as the equation with the best model fit, due to the fact that the C_p value presents the same number of predictor variables plus the constant (7.000). Such results may be directly related to the fact that there is no multicollinearity among the predictor variables ($VIF = 1$). According to Tariq et al. (2012), one of the most effective methods to solve the multicollinearity problem in multiple regression analysis is to use the factor loadings from factor analysis, through the latent variables.

Table 8. Prediction equations for tissue composition according to stepwise multiple regression analysis of the latent variables.

Dependent variable	Number	Equation	C(p)	R ²	MSE	P Value
Muscle proportion, %	1	$Y = 68.142 - 2.328 \times \text{Factor3}$	48.630	0.360	9.745	<.0001
	2	$Y = 68.142 - 1.200 \times \text{Factor1} - 2.328 \times \text{Factor3}$	31.978	0.456	8.394	<.0001
	3	$Y = 68.142 - 1.200 \times \text{Factor1} + 1.138 \times \text{Factor2} - 2.328 \times \text{Factor3}$	17.213	0.542	7.159	<.0001
	4	$Y = 68.142 - 1.200 \times \text{Factor1} + 1.138 \times \text{Factor2} - 2.328 \times \text{Factor3} - 0.795 \times \text{Factor5}$	11.027	0.584	6.589	<.0001
	5	$Y = 68.142 - 1.200 \times \text{Factor1} + 1.138 \times \text{Factor2} - 2.328 \times \text{Factor3} - 0.795 \times \text{Factor5} - 0.581 \times \text{Factor6}$	8.651	0.607	6.317	<.0001
	6	$Y = 68.142 - 1.200 \times \text{Factor1} + 1.138 \times \text{Factor2} - 2.328 \times \text{Factor3} + 0.531 \times \text{Factor4} - 0.795 \times \text{Factor5} - 0.581 \times \text{Factor6}$	7.000	0.625	6.099	<.0001
Bone proportion, %	1	$Y = 18.870 - 0.746 \times \text{Factor5}$	8.629	0.109	4.590	0.0027
	2	$Y = 18.870 - 0.673 \times \text{Factor4} - 0.746 \times \text{Factor5}$	2.181	0.198	4.186	0.0002
Total fat proportion, %	1	$Y = 9.359 + 1.518 \times \text{Factor3}$	32.048	0.283	5.907	<.0001
	2	$Y = 9.359 + 0.880 \times \text{Factor1} + 1.518 \times \text{Factor3}$	19.703	0.378	5.189	<.0001
	3	$Y = 9.359 + 0.880 \times \text{Factor1} + 1.518 \times \text{Factor3} + 0.709 \times \text{Factor5}$	12.392	0.440	4.735	<.0001
	4	$Y = 9.359 + 0.880 \times \text{Factor1} - 0.677 \times \text{Factor2} + 1.518 \times \text{Factor3} + 0.709 \times \text{Factor5}$	5.899	0.496	4.315	<.0001

Mallows' Cp (Cp); Coefficient of determination (R²); Mean square error (MSE); Croup width (CW); Carcass compactness index (CCI); Loin (LOI); Shoulder (SHO); Neck (NEC); Leg length (LL); Thorax perimeter (TP); Croup perimeter (CP); Cold carcass yield (CCY); Neck yield (NECY).

The latent variables Factor 4 and Factor 5 were included in the equation to predict bone proportion (Table 8), however, these two latent variables explained only 19.8%, which is a value expressly lower than that found in Table 5 (35.1%) with the variables LL, CCI and NEC, which were selected by the stepwise procedure with the best model developed for bone proportion. Based on the factor loadings presented in Table 7 for the latent variables corresponding to Factor 1 and 3, the variables that contributed the most were those related to body weight and measures (CEL, CIL, LL and CW) and yield characteristics (except for LEGY and LOIY). CCI, which was considered in the previous model (Table 5), was not considered in the present model since it did not show expressive factor loading in Factors 1 and 3.

According to the results presented in Table 8, four latent variables (Factors 1, 2, 3 and 5) were determined to predict the proportion of total fat, explaining 49.6% of the proportion of total fat. As observed for muscle and bone proportion, the proportion of total fat also showed lower explanation when compared to the previously proposed equation in Table 5, which showed 57.9% explanation.

Previous studies have been carried out using the factor scores from factor analysis (latent variables) in the multiple regression equations (Çelik et al., 2018; Önk et al., 2018; Tahtali, 2019; Tariq et al., 2012). The coefficients of determination found in their study ranged from 0.754 to 0.966, showing to be suitable for use as independent variables in regression analysis. In the present study, the coefficients of determination showed low to moderate values (R² ranging from 0.198 to 0.607) for tissue composition. With these results it can be concluded that the developed predictive models are not effective in predicting the composition of carcass tissues.

However, for comparison purposes, it can be concluded that the developed models of the latent variables promote a better fit in the equations based on the Cp values. Despite presenting lower R² values, the presence of moderate multicollinearity among the original variables was clear ($1 < \text{VIF} < 5$), indicating increase in the variance of the regression coefficients due to the correlations among the predictors of the model.

Conclusion

The use of latent variables from factor analysis in multiple regression models eliminates the problem of multicollinearity of the explanatory variables, thus improving the accuracy of interpretation of results by proposing a better fit of the mathematical model.

However, the R² values were moderate for muscle proportion and total fat, and low for bone proportion, which indicate that more appropriate independent variables should be used to better predict the proportion of tissues in Santa Inês sheep.

Acknowledgments

This study was financed in part by the Coordenação de Aperfeiçoamento de Pessoal de Nível Superior – Brasil (CAPES) – Finance Code 001.

References

- Alves, M. F., Lotufo, A. D. P., & Lopes, M. L. M. (2013). Seleção de variáveis stepwise aplicadas em redes neurais artificiais para previsão de demanda de cargas elétricas. *Proceeding Series of the Brazilian Society of Applied and Computational Mathematics*, 1-6. DOI: <https://doi.org/10.5540/03.2013.001.01.0144>
- Brasil. (2000). *Instrução Normativa nº 21, de 31 de Julho de 2000. Aprovar os Regulamentos Técnicos de Identidade e Qualidade de Patê, de Bacon ou Barriga Defumada e de Lombo Suíno*. Brasília, DF: Ministério da Agricultura, Pecuária e Abastecimento Secretaria de Defesa.
- Brown, A., & Williams, D. (1979). *Sheep carcass evaluation: measurement of composition using a standardised butchery method* (Memorandum, 38). Langford: Agricultural Research Council, Meat Research Council.
- Cadavez, V. A. P. (2009). Prediction of lean meat proportion of lamb carcasses. *Archiva Zootechnica*, 12(4), 46-58.
- Cardoso, D. B., Medeiros, G. R., Guim, A., Azevedo, P. S., Suassuna, J. M. A., Lima Júnior, D. M., ... Carvalho, F. F. R. (2021). Growth performance, carcass traits and meat quality of lambs fed with increasing levels of spineless cactus. *Animal Feed Science and Technology*, 272, 114788. DOI: <https://doi.org/10.1016/j.anifeedsci.2020.114788>
- Carrasco, S., Ripoll, G., Panea, B., Álvarez-Rodríguez, J., & Joy, M. (2009). Carcass tissue composition in light lambs: Influence of feeding system and prediction equations. *Livestock Science*, 126(1-3), 112-121. DOI: <https://doi.org/10.1016/j.livsci.2009.06.006>
- Çelik, Ş., Şengül, T., Söğüt, B., Inci, H., Şengül, A. Y., Kayaokay, A., & Ayaşan, T. (2018). Analysis of variables affecting carcass weight of white Turkeys by regression analysis based on factor analysis scores and ridge regression. *Revista Brasileira de Ciencia Avicola*, 20(2), 273-280. DOI: <https://doi.org/10.1590/1806-9061-2017-0574>
- Cezar, M. F., & Souza, W. H. (2007). *Carcaças ovinas e caprinas: obtenção, avaliação e tipificação*. Campina Grande, PB: UFCG.
- Daskiran, I., Keskin, S., & Bingol, M. (2017). Usability of the factor analysis scores in multiple linear regression analyses for the prediction of daily milk yield in Norduz Goats. *Journal of Agricultural Science and Technology*, 19, 1507-1515.
- Díaz, M. T., Cañeque, V., Lauzurica, S., Velasco, S., Ruíz de Huidobro, F., & Pérez, C. (2004). Prediction of suckling lamb carcass composition from objective and subjective carcass measurements. *Meat Science*, 66(4), 895-902. DOI: <https://doi.org/10.1016/j.meatsci.2003.08.013>
- Ekiz, B., Baygul, O., Yalcintan, H., & Ozcan, M. (2020). Comparison of the decision tree, artificial neural network and multiple regression methods for prediction of carcass tissues composition of goat kids. *Meat Science*, 161, 108011. DOI: <https://doi.org/10.1016/j.meatsci.2019.108011>
- Fernandes, J., Pereira Filho, J., Menezes, D., Caldas, A. C., Cavalcante, I., Oliveira, J., ... Bezerra, L. (2021). Carcass and meat quality in lambs receiving natural tannins from Mimosa tenuiflora hay. *Small Ruminant Research*, 198. DOI: <https://doi.org/10.1016/j.smallrumres.2021.106362>
- Fernandes, M. H. M. R., Resende, K. T., Tedeschi, L. O., Fernandes, J. S., Teixeira, I. A. M. A., Carstens, G. E., & Berchielli, T. T. (2008). Predicting the chemical composition of the body and the carcass of 3/4Boer × 1/4Saanen kids using body components. *Small Ruminant Research*, 75, 90-98. DOI: <https://doi.org/10.1016/j.smallrumres.2007.09.005>
- Gomes, H. F. B., Gonçalves, H. C., Neto, A. P., Cañizares, G. I. L., Roça, R. O., Marques, R. O., & Oliveira, G. M. (2013). Common factors method to predict the carcass composition tissue in kid goats. *Revista Brasileira de Zootecnia*, 42(3), 193-203. DOI: <https://doi.org/10.1590/S1516-35982013000300007>
- Hair Jr, J., Black, W., Babin, B., Anderson, R., & Tatham, R. (2009). *Análise Multivariada de Dados* (6a ed.). Porto Alegre, RS: Bookman. DOI: <https://doi.org/0-13-032929-0>
- Hankins, O. G., & Howe, P. E. (1946). *Estimation of the composition of beef carcasses and cuts* (Technical Bulletin, n. 926). U.S. Department of Agriculture. DOI: <https://doi.org/10.22004/ag.econ.169985>
- Kaiser, H. F. (1974). An index of factorial simplicity. *Psychometrika*, 39, 31-36. DOI: <https://doi.org/10.1007/BF02291575>
- Kazemi, A., Mohamed, A., Shareef, H., & Zayandehroodi, H. (2013). Optimal power quality monitor placement using genetic algorithm and Mallow's Cp. *International Journal of Electrical Power and Energy Systems*, 53, 564-575. DOI: <https://doi.org/10.1016/j.ijepes.2013.05.026>

- Lambe, N. R., Navajas, E. A., Bünger, L., Fisher, A. V., Roehe, R., & Simm, G. (2009). Prediction of lamb carcass composition and meat quality using combinations of post-mortem measurements. *Meat Science*, *81*(4), 711-719. DOI: <https://doi.org/10.1016/j.meatsci.2008.10.025>
- Lima Júnior, D. M., Carvalho, F. F. R., Maciel, M. I. S., Urbano, S. A., Oliveira, J. P. F., & Rangel, A. H. D. N. (2017). Leg tissue composition and physico-chemical parameters of sheep meat fed annatto coproduct. *Semina: Ciências Agrárias*, *38*(5), 3377-3382. DOI: <https://doi.org/10.5433/1679-0359.2017v38n5p3377>
- Macena, E. C. B. da C., Costa, R. G., Sousa, W. H. de, Cartaxo, F. Q., Ribeiro, N. L., Arandas, J. K. G., & Ribeiro, M. N. (2022). Multivariate modelling to estimate carcass characteristics and commercial cuts of Boer goats. *The Journal of Agricultural Science*, *160*(5), 371-379. DOI: <https://doi.org/10.1017/S002185962200020X>
- Mallows, C. L. (2000). Some Comments on C p. *Technometrics*, *42*, 87-94. DOI: <https://doi.org/10.1080/00401706.2000.10485984>
- Marcondes, M. I., Tedeschi, L. O., Valadares Filho, S. C., & Chizzotti, M. L. (2012). Prediction of physical and chemical body compositions of purebred and crossbred Nellore cattle using the composition of a rib section. *Journal of Animal Science*, *90*(4), 1280-1290. DOI: <https://doi.org/10.2527/jas.2011-3839>
- Miyashiro, R., & Takano, Y. (2015). Subset selection by Mallows' Cp: A mixed integer programming approach. *Expert Systems with Applications*, *42*, 325-331. DOI: <https://doi.org/10.1016/j.eswa.2014.07.056>
- Önk, K., Sari, M., & Gürcan, İ. S. (2018). Estimation of live weights at the beginning and the end of grazing season in Tuj lambs via scores of factor analysis. *Ankara Üniversitesi Veteriner Fakültesi Dergisi*, *65*(3), 261-266. DOI: https://doi.org/10.1501/Vetfak_0000002855
- Ribeiro, F. R. B., & Tedeschi, L. O. (2012). Using real-time ultrasound and carcass measurements to estimate total internal fat in beef cattle over different breed types and managements. *Journal of Animal Science*, *90*(9), 3259-3265. DOI: <https://doi.org/10.2527/jas.2011-4697>
- Santos, V. A. C., Silvestre, A. M., Azevedo, J. M. T., & Silva, S. R. (2017). Estimation of carcass composition of goat kids from joint dissection and conformation measurements. *Italian Journal of Animal Science*, *16*(4), 659-665. DOI: <https://doi.org/10.1080/1828051X.2017.1321472>
- Senra, L. F. A. de C., Nanci, L. C., Mello, J. C. C. B. S. de, Meza, & L. A. (2007). Estudo sobre métodos de seleção de variáveis em DEA. *Pesquisa Operacional*, *27*(2), 191-207. DOI: <https://doi.org/10.1590/S0101-74382007000200001>
- Tahtali, Y. (2019). Use of factor scores in multiple regression analysis for estimation of body weight by certain body measurements in Romanov Lambs. *PeerJ*, *7*(e7434), 1-11. DOI: <https://doi.org/10.7717/peerj.7434>
- Tariq, M. M., Eyduran, E., Bajwa, M. A., Waheed, A., Iqbal, F., & Javed, Y. (2012). Prediction of body weight from testicular and morphological characteristics in indigenous mengali sheep of Pakistan using factor analysis scores in multiple linear regression analysis. *International Journal of Agriculture and Biology*, *14*(4), 590-594.