Simple and robust model to estimate live weight of Ethiopian Menz sheep

Jane Wamatu[a,*], Ashraf Alkhtib[ab], Barbara Rischkowsky[a]

[a]International Center for Agricultural Research in Dry Areas (ICARDA), PO Box 5689, Addis Ababa, Ethiopia.

[b]Nottingham Trent University, school of animal, rural and environmental sciences, P.O Box NG25 0QF, Brackenhurst Lane, Southwell, Nottinghamshire, the UK.

[*]Corresponding author. Email: j.wamatu@cgiar.org (J. Wamatu)

Abstract. Heart girth (HG) bands have been predominantly used in Ethiopia by smallholder farmers, traders and extension workers to estimate live weight (LW) of livestock. They are produced using recommended and published predictive models from Ethiopia. More recently, some farmers and traders have abandoned the bands due to perceived inaccuracy of LW estimation and reverted to eye ball estimations. This study generated a novel algorithm using multiple criteria to develop a robust predictive model for LW estimation of Ethiopian Menz sheep using HG. Subsequently, recommended models currently in use in Ethiopia were evaluated for accuracy in predicting LW using data of this study. Live weight and HG of 420 Menz sheep were measured. Simple linear model (SLM), Box-Cox (SLM with $LW^{0.75}$), quadratic and allometric models were used to describe the relationship between LW and HG. Algorithms used to validate the models included data exploration, model construction and model redeployment. Results revealed that all models had similar R^2 (≈0.82). All models fitted the criteria of residuals analysis and robustness against extreme values. However, only Box-Cox was robust against data redeployment with 95[th] percentile of prediction error (PE) less than 10%. Accordingly, a Box-Cox model ($LW^{0.75} = -9.71 + 0.289(HG)$) is robust and can be used to accurately predict LW of Menz sheep. The 95th percentile of PE of existing, recommended models was higher than 10, thus they cannot be recommended to accurately predict LW of Menz sheep. This study concludes that an approach based on regressing LW on HG then selecting models with highest R^2 is inadequate to generate accurate and robust prediction models. This highlights the importance of model redeployment to generate accurate prediction models. Calibrated HG bands are suitable alternatives to weighing scales in rural areas of Ethiopia because they are cheaper and not subject to maintenance. Thus, their accuracy and robustness in estimation of LW is vital for sustainable use.

Keywords: Live weight prediction; Menz sheep; heart girth; prediction error; linear; nonlinear.

Introduction

The sub-alpine highlands of Amhara region of Ethiopia are characterized by extreme cold and frosty climate, rugged terrain, degraded soil and unreliable crop production. Livestock, particularly sheep production is the mainstay of farmers' livelihoods. Sheep represent a major source of income for smallholder farmers in this region contributing to approximately 45% of their cash income (Gizaw et al. 2012). Menz breed is one of the primary Ethiopian sheep breeds totaling over 1.5 million. The breed is concentrated in the central highlands between 2500 m and 3000 m above sea level, 39 - 40" E longitude and 10 -11" N latitude. It is mainly reared on small peasant farms in flocks of 11 (range 1-32) animals. Menz sheep are a fat-tailed hair breed of small body size with an average height of 64±1 cm (Galal 1980). Average live weight (LW) is 19.7±0.4, 33.5±1.3 kg and 38.2±0.8 kg for 6-month old lambs, yearlings and mature ewes respectively (Galal 1980). They have semi-open fleece of conical locks, coarse hair that may be 15-20 cm long and a wooly undercoat of 5-8 cm especially in the colder highlands (Galal, 1983). Meat and fiber of Menz sheep are in high demand in Ethiopia. Production and marketing of Menz sheep is, therefore, of paramount importance. Debre Berhan Agricultural Research Center (DBARC), located in Amhara, is the center of excellence for sheep production in Ethiopia and hosts large flocks of Menz sheep. This center is working with the International Center for Agricultural Research in Dry Areas (ICARDA) to improve production and productivity of the Menz breed. Efforts are being undertaken to enhance the capacity of sheep farmers in the region.

Live weight and LW change of sheep are important reflections of nutrition, management, breeding and husbandry. They are vital as indicator of growth, feed conversion efficiency (Veerkamp 1998), readiness for marketing or slaughtering (Sawyer *et al.* 1991) and dosing of drugs (Machila *et al.* 2008). Conventional weighing scales are the key standard to determine LW of sheep, provided the scales are well calibrated. However, in rural areas of Ethiopia, weighing scales are rarely used by small holder farmers due to their high costs and high demand of labor and time. Moreover, the bias in LW estimation using calibrated scales is high because their springs permanently stretch with repeated or out-of-bounds use resulting in biased measurements (Machila *et al.* 2008). Scale calibration and maintenance requires skilled technicians who are rarely found in rural areas. Heart girth (HG) has been repeatedly demonstrated to be the most useful and robust proxy for the use of scales in LW estimation of sheep (Sowande and Sobola 2008; Atta and El Khidir 2004). Heart girth bands have been predominantly used in Ethiopia by smallholder farmers, traders and extension workers to estimate LW of livestock. They are produced using recommended and published predictive models of Ethiopia (Table 1). However, more recently, some farmers and traders have abandoned the bands due to their perceived inaccuracy of LW estimation and reverted to eye ball estimations. However, some studies have demonstrated that visual LW estimation of sheep, as an alternative to using scales, lacks accuracy and is prone to error (Machila *et al.* 2008). Inaccurate LW estimates in the region has led of mistrust between farmers and market traders over selling price, which are based on LW estimates, and between farmers and extension workers over perceived failure to give proper recommendations for dosage of livestock drugs and supplementary feeds. Heart girth calibrated weight bands are usually produced either site specifically or breed specifically. In Ethiopia, predictive models of LW based on HG that are used to produce HG bands for use by farmers, traders and extension agents have been reported for Menz sheep (Getachew et al. 2008, R^2=0.83) and a mixture of Ethiopian highland sheep (Tadesse and Gebremariam 2010, R^2 =0.69; Berhe 2017, R^2=0.9). The models were generated by regressing LW on HG, and recommending models with maximum R^2. However, R^2 as a single criterion is not enough to validate models because it does not provide information about the degree to which values predicted by a model diverge from measured values (Goopy *et al.*, 2017). Furthermore, models should prove robustness in predicting other datasets other than being robust only in predicting the original data. The study sought to determine whether perceived sentiments by smallholder farmers, traders and extension agents in rural areas of Amhara region in Ethiopia that the HG bands on the market give inaccurate estimates of LW were valid. A novel algorithm, involving several criteria, was developed to provide a robust predictive equation to estimate LW in Ethiopian smallholder Menz sheep. Measures used to assess this novel algorithm included R^2, analysis of residuals (normality and homogeneity), prediction error of models and robustness of coefficients of models against bootstrapping. In addition, recommended models of Ethiopia were assessed for accuracy using data of this study.

Materials and methods

Study area

The study was conducted at DBARC Ethiopia. The station is located 120 km north-east of Addis Ababa at an altitude of 2780 m in the central highlands of Ethiopia in Amhara region (Gizaw et al. 2012).

Measurements for model development

Measurements of LW and HG for 420 recently fleeced Menz sheep (346 females and 74 males with age range of 11 to 96 months were undertaken after overnight fasting at DBARC. Live weight was measured gravimetrically using a portable spring-dial hoist scale (Camry, NTB, Camry company, China), with capacity of 100 kg and precision of 0.5 kg. The scale was calibrated using standard weights, after which 10 sheep were weighed in 3 replications to confirm reliability of LW measurements. The scale was further calibrated at 50-sheep measurement intervals. Heart girth was measured as body circumference immediately behind the front shoulder at the fourth ribs, posterior to the front leg, using an ordinary measuring tape held with 1kg tension using a light spring balance. Pregnant sheep and sick sheep as per research center records were excluded from the study.

Analytical approach

Construction of data included three main steps: Data exploration, model construction and model redeployment. In the data exploration step, data was analyzed for accuracy of collection, need for power transformation and normal distribution. In model construction, linear and nonlinear models were constructed and validated using different criteria. The third step involved redeployment of constructed models to a new data set. Three published models were validated by redeploying them to data from this study.

2.3.1. Data exploration

The accuracy of the scale in measuring LW of sheep may decline due to successive measurements of heavy sheep, therefore, the relationship between LW and the serial number of sheep was visually presented to depict the distribution of LW across the measurement process. The probability distribution of LW and HG was identified using the normal Q-Q plot. Box-Cox analysis was used to confirm whether a power transformation of LW would increase R^2 of models. Optimum power of transformation of LW was identified using a likelihood maximized Box-Cox transformation ($h(y, l) = (y_l - 1)/l$, $l = 0$; boundaries of –3 and +3 and a step of 0.25) (Box and Cox 1964). The R^2 and log likelihood values of λ value were used to identify the best power of transformation.

2.3.2. Model construction

Four models for predicting LW through HG were tested. The first model was a simple linear regression model (SLM; model 1), the second was a SLM with $LW^{0.75}$ (Box-cox, model 2), the third model was a quadratic model (QUADM; model 3) and the fourth was an allometric model (ALM; model 4). The four models are presented below:

SLM:

$$LW = a + b(HG) \quad (1)$$

Box-Cox

$$LW^{0.75} = a + b(HG) \quad (2)$$

QUADM:

$$LW = a + b(HG) + c(HG)^2 \quad (3)$$

ALM:

$$LW = a(HG)^b \quad (4)$$

Aggregated data (males and females of varying ages) using HG as a single predictor was used to construct the models. When analysis produced models that explained sufficient variation in LW, no drill-down analysis (such as disaggregating data based on gender) was carried out. Three published models that were validated included:

Getachew et al. (2008) for Menz:

$$LW = -23.4 + 0.67(HG) \quad (5)$$

Tadesse and Gebremariam (2010) for highland sheep in Ethiopia:

$$LW = -15.7 + 0.56(HG) \quad (6)$$

Berhe (2017) for highland sheep in Ethiopia:

LW= = -18.7 + 0.6(HG)..(7)

The algorithm used to validate all models for accuracy of predicting LW using HG contained various steps. These included coefficient of determination (R^2) of the models, coefficient of variation ($CV = 100 \times \frac{\sqrt{MSR}}{mean}$) of the models, where MSR is the mean squares of the residuals and mean is LW mean, Cook's distance, bootstrapping technique and analysis of residuals. Cook's distances were calculated for models 1, 2, 3 and 4 to assess existence of outliers which may have exerted a significant effect on coefficients of the models. Values of Cook's distance were compared to 50th percentile values on the F distribution ($F_{(0.5, 2, 418)} = 0.79$) (ReliaSoft 2015). Observations equal or higher than 0.79 were considered influential. Robustness of coefficients of the models was assessed using bootstrapping technique (Wood 2004). This technique involved generating 1000 bootstrap resamples (n=420 each) from the original data by random sampling with replacement. These resamples were analyzed individually and variation among resulting estimates of models (1 to 4) expressed as 95% confidence intervals. Coefficients of each model (1 to 7) were used to calculate expected LW using HG. Thereafter, residuals were calculated and standardized. Standardized residuals of each model were plotted against the serial number of sheep to identify existence of a drift in residuals. Additionally, residuals of each model were examined for normality using a normal Q-Q plot. The association between residuals and LW and HG in each model was visualized by plotting residuals against HG and LW. The 70th, 90th and 95th percentile of PE were calculated for each model as follows:

$$PE = 100 \times \left| \frac{LW_p - LW_m}{LW_m} \right|$$

where LW_p and LW_m were predicted and LW measured.

2.3.3. Model redeployment

Constructed and published models underwent a redeployment step. Bootstrap resamples were analyzed individually and variation among resulting estimates of models (a, b and c) were expressed as 95% confidence intervals. Using resampled data, predicted LW was calculated using coefficients of each model, Residuals and PE were generated. Data was analyzed using the Statistical Analysis System (SAS 2012).

Results

Data exploration

Live weight and HG of sheep ranged from 14 kg to 36 kg and 58 to 82 respectively. Fig. 1a illustrated there was no systematic relation between LW and sheep serial number. A visual inspection of normal probability plots showed that the distribution of HG and LW was close to normal (Fig. 1b). Results of Box-Cox transformation procedure showed that R^2 of λ values ranged from 0.66 to 0.82. The log likelihood values of λ ranged from -573 to -273 (Table 2). λ value which ranged from 0.25 and 1.25 had the highest R^2 values (0.82), however, λ with a value of 0.75 had both the highest R^2 (0.82) and the highest log likelihood value (-274).

Model construction

Table 3 showed that coefficients of all models (1 to 4) were significantly different from 0 (P<0.001). Coefficients (a, b and c) and corresponding standard errors of the three models are presented in Table 4. Using the coefficients (a, b and c) presented in Table 4, models 1, 2, 3 and 4 were constructed as follows:

SLM: $LW = -36.6 + 0.882(HG)$ ....(1)

Box-Cox: $LW^{0.75} = -9.71 + 0.289(HG)$ ....(2)

QUADM: $LW = -47.8 + 1.02(HG) - 0.002(HG)^2$ ....(3)

ALM: $LW = 0.001(HG)^{2.46}$ .... (4)

Standard error of (a) and (b) for SLM as a percentage of the estimate was 4% and 2% respectively. Standard error of (a), (b) and (c) for QUADM, as percentage of the estimate, was 33%, 45% and 150% respectively. The standard error of (a) and (b) for ALM, as a percentage of the estimate, was <1% and 2.5 % respectively. Cook's distance in all models were less than 0.79 (Table 3), therefore, it was not plotted against LW. All models had similar R^2 ranging from 0.814 to 0.819 (Table 3). The simple linear model had the lowest CV followed by QUADM and ALM, which were higher than SLM by 2.35 and 2.49 units respectively (Table 3). Visual inspection of normal Q-Q plots (Fig. 3a) showed that residuals of all models were almost normally distributed. Standardized residuals of models versus serial number of sheep showed that residuals of all models were scattered across serial numbers without any systematic pattern (Fig. 3b). Visual inspection of residuals versus LW plots (Fig 3c, d) showed that there was no linear relationship nor clear trends between the residuals and HG and LW in models 1, 2 and 3. Fig. 4 showed that correlations between standardized residuals of models 5, 6, and 7 and LW were negative and very strong ($r>0.79$; $P<0.001$). The PE of the 75th, 90th and 95th percentiles of Box-Cox was approximately three times less than that of our constructed models.

Model redeployment

Table 4 shows 95% confidence interval based on 1000 bootstrapping resamples. The 95% confidence interval of SLM estimates were 5.3 units for (a) and 0.075 units for (b). The 95% confidence interval estimates of Box-Cox model were 0.15 units for (a) and 0.002 units for (b). The 95% confidence intervals of QUADM estimates was 52.4 units for (a), 1.54 units for (b) and 0.005 units for (c). The 95% confidence interval of ALM estimates was ~0.001 units for (a) and 0.18 units for (b). Table 5 shows percentiles of PE of constructed and published models based on 1000 resamples. Out of all constructed and published models, only Box-Cox model had PE percentiles less than ~10.

Discussion

Data exploration

An examination of normal Q-Q plot showed that observed values of LW and HG were close to predicted values and the distribution of points around trend lines was symmetric. That suggests the distribution of LW and HG was close to normal with slight deviation. Thus, transforming LW to decrease PE and increase R^2 of the prediction model might be required (Lesosky *et al.* 2013). This result is confirmed by results of Box-Cox procedure which showed that λ with a value of 0.75 had highest R^2 and highest log likelihood values. Thus a power transformation (0.75) might increase the accuracy of LW prediction by confirming the normality of LW (McDonald 2009). Accordingly, SLM with transformed LW was constructed. Values of Cook's distance for all sheep were less than the critical value (0.79) which means there were no outliers in the data. Data exploration step confirmed that all sheep in this study should be included in the model construction step and SLM with transformed LW should be constructed in addition to SLM, QUAD and ALM.

Model construction

All constructed models had high R^2 (R^2> 0.8), however, the coefficient of variation of SLM and Box-Cox was considerably less than CV of QUADM and ALM. That suggests that the bias in QUADM and ALM prediction was higher than that of SLM. Accordingly, deeper analysis of residuals is required to choose the best-fit model. Observed values of residuals were close to expected normal values suggesting that residuals were almost normally distributed. Residuals of the constructed models were symmetrically distributed around 0, indicating that residuals of the constructed models were not biased to a positive nor negative tail. Weak correlation between residuals and serial number of sheep in all models confirms that accuracy of models was constant alongside the measurement process. There was absence of any systematic relation between residuals and LW and HG for all models, suggesting accuracy of predicting LW of Menz sheep using HG was constant for all sheep regardless of their LW or HG. Accordingly, prediction of LW by constructed models was equal for all sheep regardless of the order in measurement process, LW and HG. All constructed models fulfilled criteria of normality and homogeneity of residuals. However, the magnitude of PE will be decisive in selecting the best-fit model among constructed models. Box-Cox model had the lowest 95th percentile of PE among constructed models. Additionally, only Box-Cox model had 95th percentile of PE less than 10, indicating that only Box-Cox model could be used to predict LW of Menz sheep for husbandry, management and veterinary purposes.

Model redeployment

The robustness of the Box-Cox model to predict LW of Menz sheep not in this study needed to be investigated. Ninety five percent of resamples' coefficients of Box-Cox model were in a very narrow range (a±1.54% for (a) and b±0.7% for (b)) compared to other constructed models which had wide range of confidence interval (<1% to 110%). Thus, Box-Cox was the only model with coefficients that were robust against bootstrapping. The 95th percentile of PE of Box-Cox was less than 10%, which is considered the critical PE for purposes of estimating live weights for veterinary, management, breeding and nutrition. Box-Cox model had homogenous and normally distributed residuals. It was robust against bootstrapping. Moreover, it predicted LW of Menz sheep with a level of precision suitable to breeding, husbandry, nutrition and veterinary services. Accordingly, Box-cox model provides the best estimate and could be accurately used to predict LW of Menz sheep.

The strong and negative correlation between residuals and LW in the published models (5, 6 and 7) means that the sheep with heavier LW had a smaller residual suggesting that the accuracy of published models in predicting LW of Menz sheep depends on LW. Furthermore, PE of all percentiles of published models (5, 6 and 7) exceeded 10% suggesting that the published models are not suitable for the estimation of LW of Menz sheep for veterinary, management, breeding and nutrition purposes. The model of Getachew *et al.* (2008) -Model 5- which was constructed to predict LW of Menz sheep was not sufficiently able to predict LW of sheep in this study. The reason could be that Getachew *et al.* (2008) used a simple approach to generate his model without considering analysis of residuals of the model as well as the model redeployment step. The model did not consider the magnitude of PE which critically affects precision of HG measurements to estimate LW. Although a PE of 20% may be acceptable for setting dosage rates for veterinary purposes, a PE of 10% or greater is problematic when using HG measurements to evaluate production-related traits such as growth and feed conversion ratio which require accurate LW determination (Leach and Roberts 1981). Getachew *et al*'s model did not examine Cook's distance which is an important criterion to determine the existence of observations which might have a significant effect on coefficients of a regression model (ReliaSoft 2015). Kmenta (1986) reported that analyzing residuals from normality, drift and homogeneity is an important criterion to validate regression models.

Transforming LW before including it in a simple linear regression was reported to decrease PE to less than 20% and to increase R^2 up to 0.98 (Lesosky *et al.* 2013). The approach used to generate the model considered only a linear model although allometric (Atta and El Khidir 2004), quadratic and exponential relationships between HG and LW (Buvanendran *et al.* 1980; Nesamvuni *et al.* 2000; Francis *et al.* 2004) have been reported. Goopy *et al.* (2017) used an algorithm to validate regression models which included R^2, the root mean squared error, Cook's distance and PE. However, additional analyses are still required for validation of models. These include identification of drift in residuals and standard error of coefficients of models. Furthermore, developed models should prove robustness in predicting successive datasets. This underpins the importance of using an appropriate analytical approach to generate models which predict LW of sheep using heart girth.

Published models 6 and 7, which were constructed using different breeds of Ethiopian highland sheep, could not predict LW of sheep of this study with PE less than 10. This might be due to variation in morphological characteristics among sheep breed which may influence the relationship between LW and HG. This affirms that prediction models of LW of sheep need to be breed-specific. Accordingly, Box-Cox model generated for Menz sheep in this study should not be generalized to other Ethiopian sheep breeds. Further studies are necessary to determine robust models for other Ethiopian sheep breeds.

Indeed, the perception of smallholder farmers, traders and extension agents in rural areas of Amhara region in Ethiopia that the currently available HG band on the market gives inaccurate estimates of LW appear valid from our studies.

Conclusion

This study underpins the importance of using appropriate analytical approaches to generate models which predict LW of sheep using heart girth.

Conflict of Interest

The authors declare that they have no conflict of interest regarding the publication of this paper.

Acknowledgements

This study is part of the International Fund for Agricultural Development (IFAD) project "Improving the performance of pro-poor value chains of sheep and goats for enhanced livelihoods, food and nutrition security in Ethiopia". The authors would like to acknowledge the IFAD funding. Data collection from Debre Berhan Research Center is appreciated.

References

Atta M, El Khidir O (2004) Use of heart girth, wither height and scapuloischial length for prediction of liveweight of Nilotic sheep. *Small Ruminant Research* **55**, 233–237.

Berhe W (2017) Relationship and prediction of body weight from morphometric traits of indigenous highland sheep In Tigray, Northern Ethiopia. *Journal of Biotechnology and Bioinformatics* **3**, 1–5.

Box G, Cox D (1964) An analysis of transformations. *Journal of the Royal Statistical Society Series B Methodological* **26**, 211–252.

Buvanendran V, Umoh J, Abubakar B (1980) An evaluation of body size as related to weight of three West African breeds of cattle in Nigeria. *The Journal of Agricultural Science* **95**, 219–224. doi:10.1017/ S0021859600029476.

Francis J, Sibanda S, Kristensen T (2004) Estimating body weight of cattle using linear body measurements. *Zimbabwe Veterinary Journal* **33**, 15–21. doi:10.4314/zvj.v33i1.5297.

Galal, E. (1983). Sheep germplasm in Ethiopia. *Animal Genetic Resources Information, 1*, 5-12. doi:10.1017/S1014233900000018

Galal, E. (1980). A contribution towards a plan for sheep and goat research and development in Ethiopia. Institute of Agricultural Research, Addis Ababa.

Getachew T, Sharma A, Tibbo M, Haile A (2008) Characterization of Menz and afar indigenous sheep breeds of smallholders and pastoralists for designing community-based breeding strategies in Ethiopia. Haramaya.

Gizaw S, Aschalew T, Lemma W, Beneberu T, Shenkute G, Wamatu J, Thorpe W, Duncan A (2012) Characterization of the farming and livestock production systems and the potentials to enhance productivity through improved feeding in the Subalpine Highlands of Amhara region, Ethiopia. (Addis Ababa)

Goopy JP, Pelster DE, Onyango A, Marshall K, Lukuyu M (2017) Simple and robust algorithms to estimate liveweight in African smallholder cattle. *Animal Production Science*. doi:10.1071/AN16577.

Kmenta J (1986) ‘Elements of Econometrics.’ (Macmillan: New York)

Leach T, Roberts C (1981) Present status of chemotherapy and chemoprophylaxis of animal trypanosomiasis in the eastern hemisphere. *Pharmacology & Therapeutics* **13**, 91–147. doi:10.1016/0163-7258(81) 90069-3.

Lesosky M, Dumas S, Conradie I, Handel I, Jennings A, Thumbi S, Toye P, BM de CB (2013) A live weight–heart girth relationship for accurate dosing of east African shorthorn zebu cattle. *Tropical Animal Health and Production* **45**, 311–316. doi:10.1007/s11250-012-0220-3.

Machila N, Fèvre E, Maudlin I, Eisler M (2008) Farmer estimation of live bodyweight of cattle: implications for veterinary drug dosing in East Africa. *Preventive Veterinary Medicine* **87**, 394–403. doi:10.1016/j. prevetmed.2008.06.001.

McDonald JH (2009) ‘Handbook of Biological Statistics.’ (Sparky House Publishing: Baltimore) doi:10.1017/CBO9781107415324.004.

Nesamvuni A, Mulaudzi J, Ramanyimi N, Taylor G (2000) Estimation of body weight in Nguni-type cattle under communal management conditions. *South African Journal of Animal Science* **30**, 97–98.

ReliaSoft (2015) 'Experiment Design & Analysis Reference.' (ReliaSoft Corporation: Tucson)

SAS (2012) 'SAS/STAT 12.1 User's Guide.' (SAS Inc: Cary)

Sawyer G, Barker D, Morris R (1991) Performance of young breeding cattle in commercial herds in the south-west of Western Australia. 2. Liveweight, body condition, timing of conception and fertility in first-calf heifers. *Australian Journal of Experimental Agriculture* **31**, 431–441. doi:10.1071/EA9910431.

Sowande O, Sobola O (2008) Body measurements of west African dwarf sheep as parameters for estimation of live weight. *Tropical Animal Health and Production* **40**, 433–439.

Tadesse A, Gebremariam T (2010) Application of linear body measurements for live body weight estimation of highland sheep in Tigray region, North-Ethiopia. *Journal of the dry lands* **3**, 203–207.

Veerkamp R (1998) Selection for economic efficiency of dairy cattle using information on live weight and feed intake: A Review1. *Journal of Dairy Science* **81**, 1109–1119. doi:10.3168/jds.S0022-0302 (98)75673-5.

Wood M (2004) Statistical inference using bootstrap confidence intervals. *Significance* **1**, 180–182. doi:10.1111/j.1740-9713.2004.00067.x.

Table 1. Summary of studies investigating the relationship between heart girth and live weight (LW) for Ethiopian sheep

R^2, coefficient of determination. LW, live weight

Reference	(Getachew *et al.* 2008)	(Tadesse and Gebremariam 2010)	(Berhe 2017)
Breed	Menz	Highland sheep	Highland sheep
n of sheep	1186	285	257
LW (kg) range	20.6±0.15	20.5±2.98	Not available
R^2	0.83	0.69	0.9
Model	LW= -23.4 + 0.67(HG)	LW= -15.7 + 0.56(HG)	LW= -18.7 + 0.6(HG)

Table 2. λ values and their corresponding coefficient of determination and log likelihood values resulting from the Box-Cox transformation procedure

R^2, coefficient of determination

λ	R^2	Log-likelihood
-3	0.66	-573
-2.75	0.68	-542
-2.5	0.72	-512
-2.25	0.721	-483
-2	0.73	-455
-1.75	0.75	-429
-1.5	0.761	-404
-1.25	0.772	-381
-1	0.783	-359
-0.75	0.792	-339
-0.5	0.8	-322
-0.25	0.81	-307
0	0.81	-295
0.25	0.82	-285
0.5	0.82	-278
0.75	0.829	-274
1	0.821	-273
1.25	0.82	-274
1.5	0.81	-277
1.75	0.81	-283
2	0.8	-292
2.25	0.822	-301
2.5	0.792	-313
2.75	0.783	-326
3	0.772	-340

Table 3. Percentiles of PE and cook's distance of models for estimating live weight of Menz sheep using heart girth

R^2, coefficient of determination. CV, coefficient of variation. PE, prediction error. ***, <0.001

Model	1 SLM	2 Box-Cox	3 QUADM	4 ALM
R^2	0.819	0.82	0.819	0.814
CV	5.51	5.96	7.86	8
P value	***	***	***	***
Percentiles of PE				
75^{th}	9.76	3.23	9.82	9.18
90^{th}	13.9	4.77	13.7	14
95^{th}	18.6	6.32	18.4	19.6
Cook's distance				
75^{th}	0.004	0.002	0.003	0.004
90^{th}	0.007	0.007	0.007	0.008
95^{th}	0.009	0.009	0.009	0.009

Table 4. Regression coefficients and bootstrapping confidence interval of estimating live weight of Menz sheep

LW, live weight (kg). HG, heart girth (cm). SE, standard error

		Bootstrap for coefficients	
		95% confidence interval	
Model	Coefficients	Lower	Upper
Model 1			
LW=a + b(HG)			
a(SE)	-36.6(1.45)	-39.4	-34.1
b(SE)	0.882(0.021)	0.846	0.921
Model 2			
$LW^{0.75}$= a + b(HG)			
a(SE)	-9.71(0.489)	-9.84	-9.69
b(SE)	0.298(0.007)	0.298	0.3
Model 3			
LW= a + b(HG) + $c(HG)^2$			
a(SE)	-47.8(15.8)	-74.4	-22
b(SE)	1.02(0.456)	0.443	1.98
c(SE)	-0.002(0.003)	-0.008	0.003
Model 4			
LW= $a(HG)^b$			
a(SE)	0.001(0.00)	0.00	0.001
b(SE)	2.46(0.061)	2.37	2.55

Table 5. Models redeployment: Percentiles of PE of constructed and published models based on data of 1000 resamples

Model	75th	90th	95th
SLM	9.69	13.6	18.5
Box-Cox	6.83	9.5	10.1
QUADM	49.1	55.1	59.2
ALM	56.4	55.6	64.4
Published			
(Getachew *et al.* 2008)	12.7	16.8	18.5
(Berhe 2017)	13.6	18.3	19.5
(Tadesse and Gebremariam 2010)	14.1	18.1	20

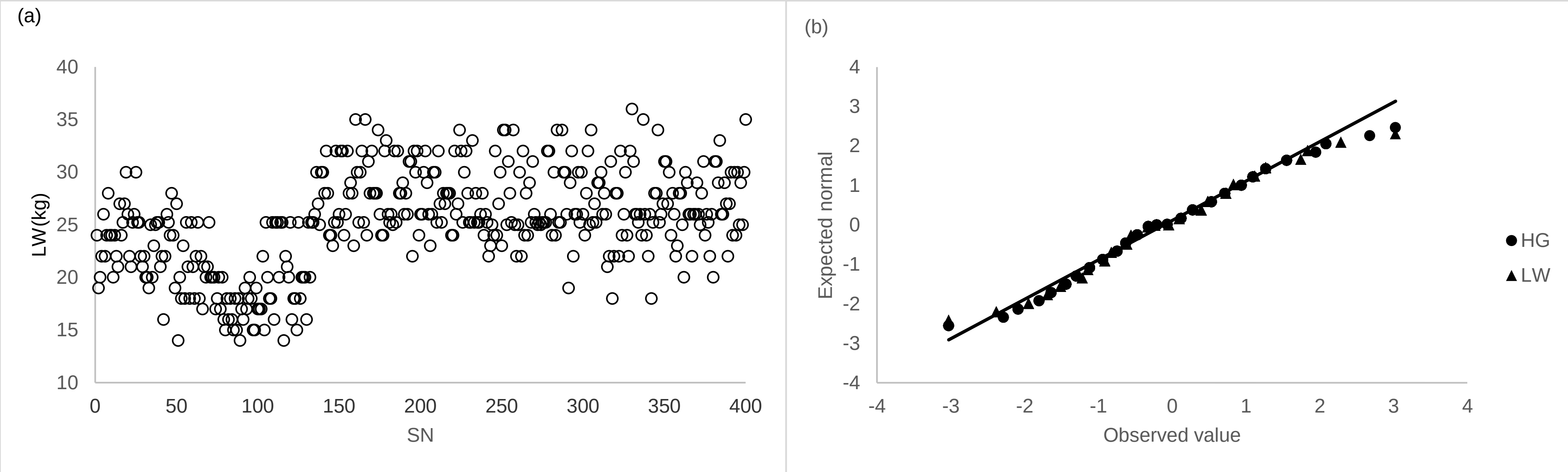

Fig. 1. Data diagnoses: (a), Q-Q normal plot of live weight and heart girth of Menz sheep; (b), Live weight vs. serial number; LW, live weight; SN, serial number.

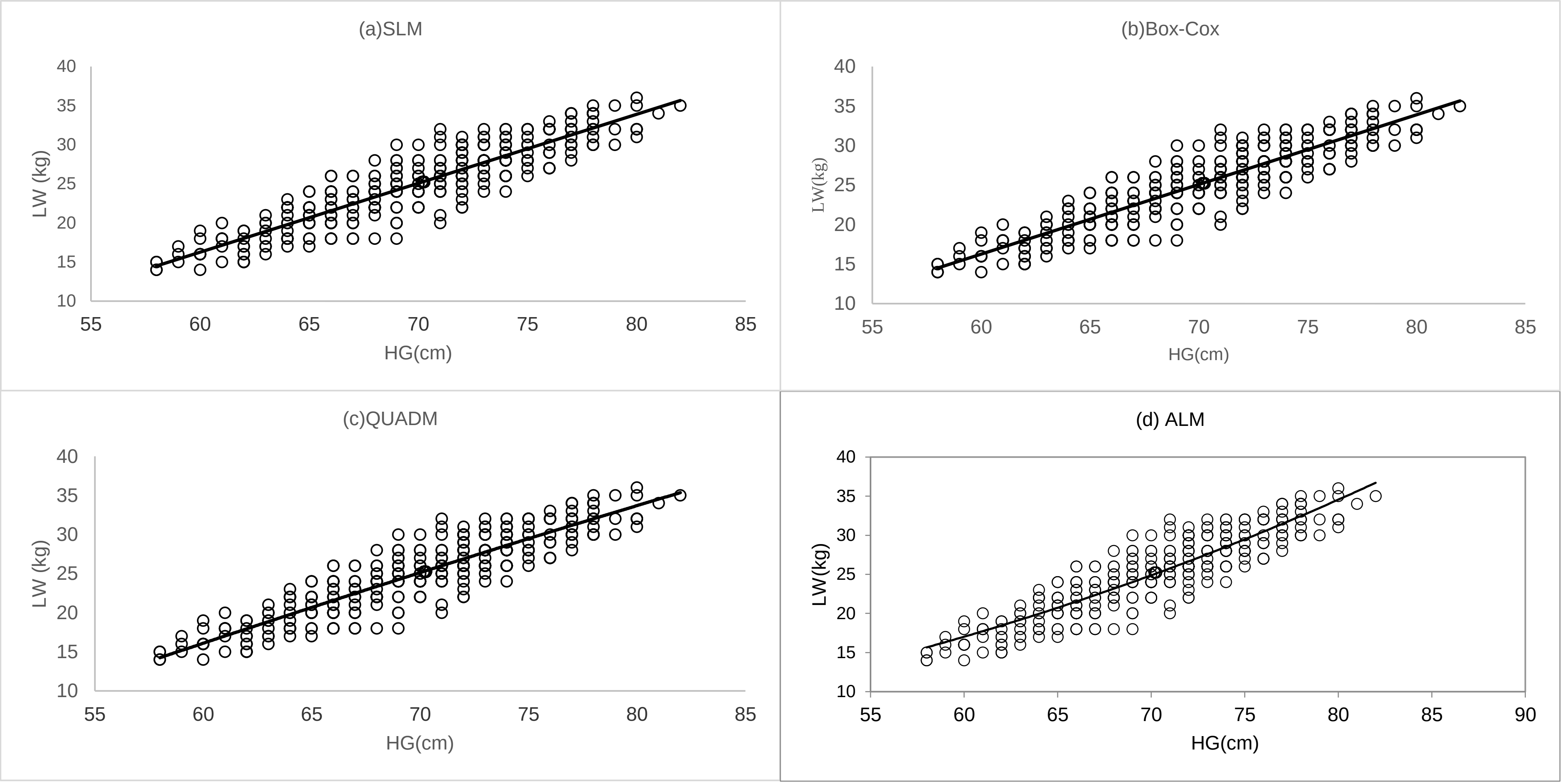

Fig. 2. Sheep live weights (LW kg) as a function of heart girth (HG cm); SLM, simple linear model; Box-Cox, simple linear model with $LW^{0.75}$, QUADM, quadratic model; ALM, allometric model.

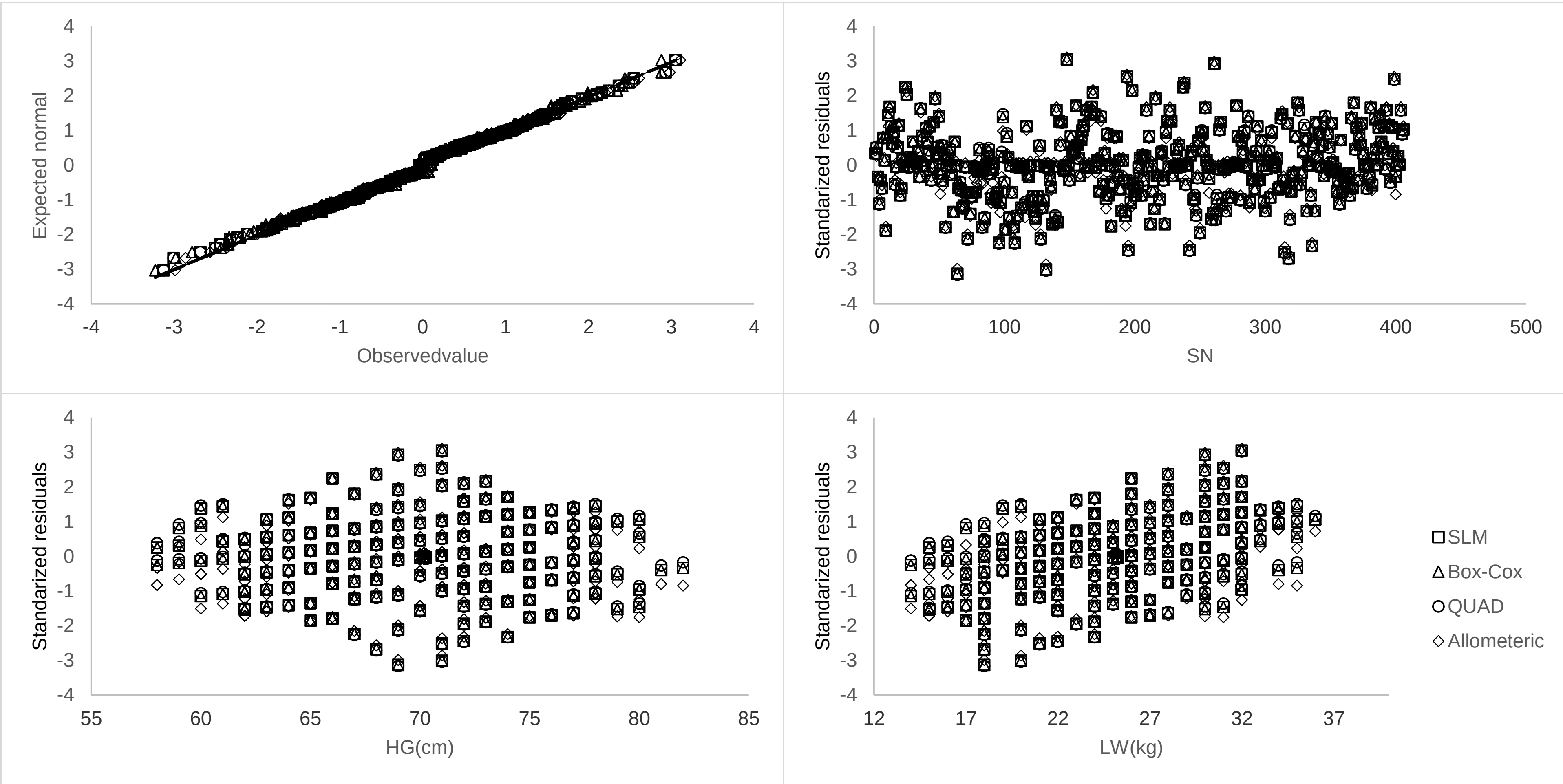

Fig. 3. Standardized residual plots for regression models (SLM, simple linear model; Box-Cox, simple linear model with $LW^{0.75}$, QUADM, quadratic model; ALM, allometric model); LW, live weight; HG, heart girth; SN, serial number.

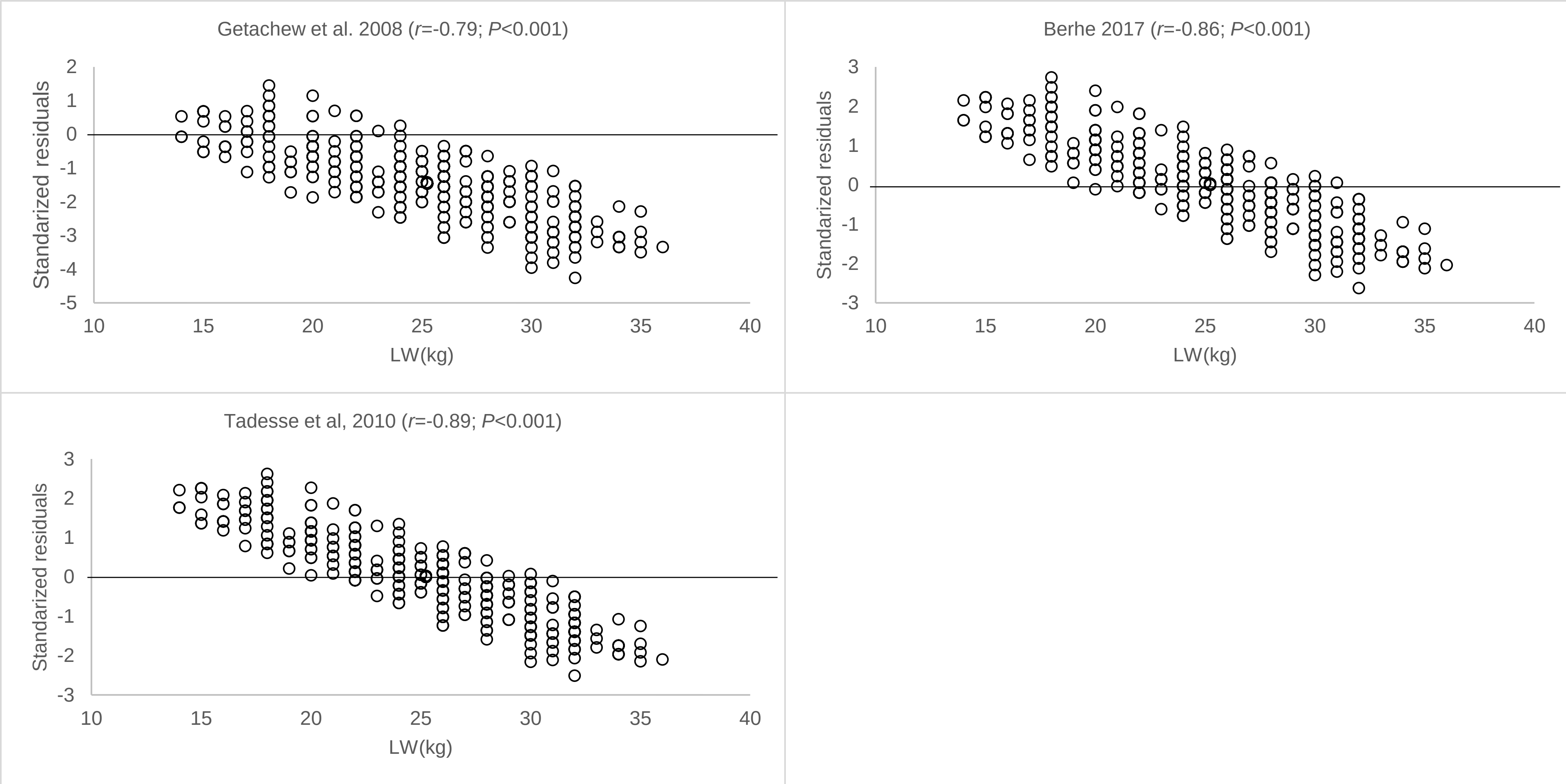

Fig. 4. Standardized residual plots for regression models (Getachew *et al.* 2008 (model 5); Tadesse and Gebremariam 2010 (model 6); Berhe 2017 (model 7)); HG, heart girth (cm); LW, live weight (kg).