Research Square

Preprints are preliminary reports that have not undergone peer review.
They should not be considered conclusive, used to inform clinical practice,
or referenced by the media as validated information.

Age-dependent morphometric variation and internally validated live-weight prediction in phenotypically identified Waila sheep of northern Cameroon

Nsiyapnze Katte Yato Katte

yatokatte@gmail.com

University of Dschang Faculty of Agronomy and Agricultural Sciences: Universite de Dschang Faculte d'Agronomique et des Sciences Agricoles https://orcid.org/0009-0005-7453-0914

Temfack Mekenjio Franck

University of Dschang Faculty of Agronomy and Agricultural Sciences: Universite de Dschang Faculte d'Agronomique et des Sciences Agricoles

Shey Ndogmi Yoniwo

University of Dschang Faculty of Agronomy and Agricultural Sciences: Universite de Dschang Faculte d'Agronomique et des Sciences Agricoles

Blaise Arnaud Hako Touko

Universite de Dschang Faculté d'Agronomique et des Sciences Agricoles: Universite de Dschang Faculte d'Agronomique et des Sciences Agricoles

Fernand Tendonkeng

Universite de Dschang Faculté d'Agronomique et des Sciences Agricoles: Universite de Dschang Faculte d'Agronomique et des Sciences Agricoles

Research Article

Keywords: body measurements, cross-validation, neck circumference, prediction equation, sheep, Waila

Posted Date: November 11th, 2026

DOI: https://doi.org/10.21203/rs.3.rs-11314169/v1

License: This work is licensed under a Creative Commons Attribution 4.0 International License.
Read Full License

Abstract

Live-weight equations require population-specific validation for use where scales are unavailable. We analysed 287 phenotypically identified Waila sheep (177 males, 110 females) from four departments of northern Cameroon. Sex differences were modelled across four age classes (≤ 12, > 12–24, > 24–36 and > 36–60 months), including sex-by-age interactions and department adjustment. Morphometric structure was assessed using principal component analysis and Ward clustering. Predictor selection was repeated within validation folds; fixed linear and machine-learning models used identical 10-fold partitions repeated ten times. Sex differences in live weight and four principal measurements depended on age (global interaction $P \leq 0.011$). Adjusted male–female weight contrasts were negligible at ≤ 12 months and ranged from + 9.21 to + 20.91 kg thereafter. One general-size component explained 69.3% of morphometric variance. Clustering this component favoured six groups by silhouette width within the tested range of two to six, partitioning an age- and sex-associated size continuum without establishing distinct morphotypes. Neck circumference had the strongest pooled correlation with weight ($r = 0.800$) and lower cross-validated root mean squared error (RMSE) than chest girth (10.08 versus 10.70 kg), but no consistent within-sex advantage. The three-measurement equation combining neck circumference, scapulo-ischial length and chest girth achieved repeated cross-validation $R^2 = 0.780$ and RMSE = 7.79 kg. The six-measurement stepwise equation achieved nested cross-validation $R^2 = 0.790$ and RMSE = 7.63 kg. Fixed machine-learning configurations showed no consistent improvement across error metrics. These internally validated equations support approximate weight estimation but require independent-flock validation before decision-specific use.

Introduction

Sheep contribute meat, cash income and readily saleable assets to smallholder systems in sub-Saharan Africa (FAO 2015). Their management and economic roles vary across Cameroon's agroecological zones (Tchouamo et al. 2005). Wilson (1991) described a large white sheep as Waïla in northern Cameroon and used Balami for a related Nigerian type; Waila is therefore used here for the phenotypically identified population sampled in Cameroon. Recent characterisation of Poulfouli sheep provides regional context (Douandji et al. 2025), but concerns a different named population. Quantitative information on Waila morphometry and locally calibrated live-weight prediction remains limited.

Linear body measurements can provide approximate live-weight estimates where livestock scales are unavailable. Chest or heart girth is frequently the leading single predictor in sheep (Thys and Hardouin 1991; Markos et al. 2023; Pelotshweu et al. 2025), although preferred measurements and coefficients differ among populations and between sexes (Baffour-Awuah et al. 2000; Avalos-Castro et al. 2023). Validation must mimic application to new animals. Selecting predictors using the complete dataset before cross-validation can inflate accuracy estimates; repeating selection within training folds reduces this bias (Varma and Simon 2006; Cawley and Talbot 2010). Machine-learning algorithms may capture nonlinearity and interactions, but should be compared with interpretable linear models on identical partitions.

This study aimed to (1) describe live weight and morphometric measurements and assess whether sex associations varied across age classes after adjustment for department; (2) characterise the dominant multivariate pattern without equating exploratory clusters with biological subpopulations; and (3) develop and internally validate live-weight equations, including field-oriented reduced models and prespecified machine-learning benchmarks. A sensitivity analysis excluded sheep aged 12 months or younger; the retained group is termed the older-than-12-month subset and not an adult subset because dental, skeletal and reproductive maturity vary among sheep (Cocquyt et al. 2005).

Materials and methods

Ethical oversight

The study protocol was approved by the Scientific and Ethical Committee of the University of Dschang. Field activities were authorised by the Ministry of Livestock, Fisheries and Animal Industries (MINEPIA) through its regional and departmental delegations. The study was conducted in accordance with the animal welfare standards for the use of animals in research set out in Chap. 7.8 (Use of animals in research and education) of the World Organisation for Animal Health (WOAH) Terrestrial Animal Health Code, and with the CIOMS–ICLAS International Guiding Principles for Biomedical Research Involving Animals (2012). Procedures were limited to non-invasive external body measurements and weighing during routine handling; no samples were taken and no animal was sedated or treated. Animal owners provided informed consent before enrolment and remained present during measurement.

Study area, animals and sampling

The cross-sectional survey was conducted from March to June 2022 in the Sudano-Sahelian zone of Cameroon, approximately 8°36'–12°54' N and 12°30'–15°42' E. Accessible villages and holdings were visited with assistance from local Zootechnical and Veterinary Centres. Animals were enrolled by convenience from smallholder flocks, household herds and fattening units; market animals were excluded. The sample comprised 287 sheep from Diamaré (n = 95), Bénoué (n = 91), Mayo-Rey (n = 52) and Mayo-Louti (n = 49) (Fig. 1). Flock identifiers were not recorded.

Trained field personnel identified sheep phenotypically as Waila using a uniform white coat, pendulous ears, convex facial profile, long frame, horns commonly present in males and absent or reduced in females, and a thin, medium-length tail. Owner accounts of local breed history were considered, and animals with evident crossbreeding were not enrolled. Breed identity was not confirmed molecularly. Age in months was estimated from owner reports and checked against dentition (Hamito 2009). Classes were defined from recorded age in months as ≤ 12, >12–24, > 24–36 and > 36–60 months. These closed and open boundaries are stated explicitly because 11 animals were recorded at exactly 12 months. Physiological state was recorded for all females as pregnant, lactating, non-pregnant and non-lactating, or nulliparous.

Measurements and data audit

Live weight was recorded using a Crane Scale digital hanging scale with a capacity of 315kg and a readability of 0.1 kg, after validating its accuracy through a span calibration procedure using certified test weights. Linear measurements were recorded to 0.1 cm in the audited dataset and were collected with a flexible measuring tape, measuring staff and a graduated ruler. Sheep stood square on level ground with the head in a natural position. One trained operator made all measurements; replicate measurements were not available. Feeding and rumen-fill status were not standardised.

Fifteen morphometric measurements were analysed (Supplementary File 1, Fig. S1). Poll-to-buttock length (PBL; *longueur du corps/longueur totale* in the field protocol) extended from the poll (chignon; occipital region) to a vertical plane tangent to the most caudal point of the buttock. It was distinct from scapulo-ischial length (SIL), measured from the point of shoulder to the point of buttock and corresponding to conventional trunk/body length. Pelvic length (PL) extended from the point of hip to the point of buttock. Height at withers (HW), height at back (HB) and height at rump (HR) were recorded vertically from the supporting surface to their respective dorsal landmarks. Chest girth (CG) was measured at the girth passage immediately behind the forelimbs; chest depth (CD) was the vertical thoracic distance at that site. Cannon-bone circumference (CC) was recorded around the middle of the anterior metacarpus. Head length (HL), ear length (EL), horn length (HoL), neck length (NL), neck circumference (NC) and tail length (TL) followed the field landmarks shown in Fig. S1. PBL was not used as conventional body length in zoometric indices.

Statistical analysis

Analyses were performed in Python 3.12.13 using NumPy, pandas, SciPy, and scikit-learn, with random seed 42. Continuous variables are presented as mean ± standard deviation (SD). Male–female differences were estimated using Welch tests and 95% confidence intervals; Holm adjustment-controlled multiplicity across the 15 measurements. Standardised differences were expressed as Cohen's d with nonparametric bootstrap confidence intervals.

For live weight, NC, HW, SIL and CG, the primary adjusted model was trait ~ sex + age class + department + sex × age class. A partial F test assessed the three interaction terms jointly, and model-based male–female contrasts were calculated within each class. To assess sensitivity to age representation, identical repeated 10 × 10-fold partitions were used to compare this model with trait ~ sex + age in months + department. Because these models are not nested, no nested-model F test was used for that comparison; paired subject-level bootstrap intervals (10,000 resamples) were calculated for differences in cross-validated RMSE and MAE.

PCA was applied to standardised non-horn measurements after Kaiser–Meyer–Olkin (KMO) and Bartlett tests. Components with eigenvalue > 1 were retained. Ward hierarchical clustering was used; retained scores for k = 2–6 were compared using average silhouette width (Rousseeuw 1987). Stability was assessed with 100 nonparametric row-bootstrap replicates. In each replicate, 287 animals were sampled

with replacement, standardisation and PCA were refitted, Ward clustering was repeated for each k, and adjusted Rand index (ARI) compared bootstrap labels with original labels indexed by sampled identifiers. Duplicates remained duplicated, omitted animals did not enter that replicate, and no label matching was required because ARI is permutation invariant (Hennig 2007).

Pearson correlations and simple regressions were used to relate live weight to each of the 14 non-horn measurements. Identical repeated folds compared NC-only and CG-only equations overall, by sex, in the older-than-12-month subset, and within sex in that subset. Paired bootstrap confidence intervals quantified NC-minus-CG differences in RMSE and MAE. An adjusted model included NC, CG, sex, age class, department, NC × sex and CG × sex.

Forward stepwise multiple regression used P-to-enter = 0.05 overall and by sex and age class; HoL, EL and TL were excluded from the candidate pool. The principal performance estimate used 10-fold cross-validation with selection repeated inside each training fold (five folds for the two smallest age classes). R^2, RMSE, mean absolute error (MAE), mean absolute percentage error (MAPE) and median absolute percentage error were calculated from held-out predictions. Fixed one-, two-, three- and six-measurement equations were compared on identical repeated 10 × 10-fold partitions. Applicable ranges were the observed minima and maxima; no equation was extrapolated beyond those ranges.

Linear regression, random forest, gradient boosting and radial support-vector regression used the same 14 non-horn predictors and identical repeated partitions. Configurations were prespecified: random forest, 500 trees and minimum leaf size 3; gradient boosting, 400 trees, depth 3 and learning rate 0.05; and support-vector regression, C = 100, ε = 0.5 and γ = scale after predictor standardisation. Calibration and Bland–Altman agreement used held-out predictions (Bland and Altman 1986). Geographic robustness used leave-one-department-out validation with stepwise selection repeated in each training set. Influence was assessed by refitting after excluding Cook's distance > 4/n; heteroscedasticity by the Breusch–Pagan test. Tests were two-sided with α = 0.05.

Results

Sample description and age-dependent sex differences

The sample contained 37 sheep aged ≤ 12 months, 103 aged > 12–24 months, 93 aged > 24–36 months and 54 aged > 36–60 months. Mean live weight was 51.31 ± 16.66 kg (range 11.6–108.4 kg). Males were heavier than females in unadjusted analysis by 12.52 kg (95% CI 9.15–15.88; Table 1).

Table 1
Selected descriptive statistics and unadjusted sex comparisons (n = 287)

Trait	Overall	Males (n = 177)	Females (n = 110)	M − F (95% CI)	Holm P
Live weight (kg)	51.31 ± 16.66	56.11 ± 17.65	43.59 ± 11.31	+ 12.52 (9.15, 15.88)	< 0.001
Poll-to-buttock length (cm)	120.61 ± 13.92	122.56 ± 14.57	117.47 ± 12.24	+ 5.09 (1.94, 8.24)	0.015
Scapulo-ischial length (cm)	72.70 ± 9.26	73.84 ± 10.19	70.88 ± 7.21	+ 2.96 (0.93, 4.98)	0.024
Chest girth (cm)	85.91 ± 12.11	87.22 ± 12.71	83.79 ± 10.80	+ 3.43 (0.66, 6.19)	0.046
Height at withers (cm)	77.63 ± 7.90	78.78 ± 8.66	75.78 ± 6.09	+ 3.00 (1.29, 4.72)	0.007
Neck circumference (cm)	32.53 ± 6.66	34.78 ± 7.01	28.91 ± 3.96	+ 5.87 (4.60, 7.15)	< 0.001
Cannon-bone circumference (cm)	10.79 ± 1.40	11.04 ± 1.47	10.40 ± 1.17	+ 0.64 (0.33, 0.95)	< 0.001

Values are mean ± SD. Complete results, coefficient-of-variation values and Cohen's d appear in Supplementary File 1, Table S3.

The sex-by-age-class interaction was significant for live weight, NC, HW, SIL and CG (global P = 0.0005, 0.005, 0.003, 0.011 and 0.002, respectively). Consequently, a single adjusted male coefficient was not used as the primary interpretation. At ≤ 12 months, adjusted male–female contrasts were small and statistically uncertain. Live-weight contrasts were + 9.21 kg at > 12–24 months, + 9.42 kg at > 24–36 months and + 20.91 kg at > 36–60 months; NC contrasts followed the same age-dependent pattern (Table 2).

Table 2
Adjusted male–female contrasts within age class

Age class (months)	LW (kg)	NC (cm)	HW (cm)	SIL (cm)	CG (cm)
≤ 12	−0.01 (− 7.27, 7.25)	+ 1.01 (− 2.26, 4.28)	−3.97 (− 7.94, 0.00)	−4.38 (− 9.04, 0.27)	−1.13 (− 7.18, 4.92)
> 12–24	+ 9.21 (4.76, 13.67)	+ 6.18 (4.18, 8.19)	+ 2.96 (0.52, 5.40)	+ 1.33 (− 1.53, 4.19)	+ 0.70 (− 3.02, 4.41)
> 24–36	+ 9.42 (4.83, 14.02)	+ 3.77 (1.70, 5.84)	+ 1.82 (− 0.69, 4.33)	+ 2.91 (− 0.04, 5.86)	−0.48 (− 4.31, 3.35)
> 36–60	+ 20.91 (14.39, 27.42)	+ 8.18 (5.25, 11.11)	+ 6.22 (2.66, 9.79)	+ 6.02 (1.84, 10.19)	+ 11.60 (6.18, 17.03)

Entries are male minus female estimate (95% CI) from trait ~ sex × age class + department. Global interaction P values were 0.0005 (LW), 0.005 (NC), 0.003 (HW), 0.011 (SIL) and 0.002 (CG).

Across identical repeated folds, the age-class interaction model reduced live-weight RMSE by 0.51 kg relative to the linear-age model, but the paired 95% CI included zero (-1.13 to 0.12 kg). For CG, the reduction was 0.61 cm (95% CI -1.10 to -0.12); intervals for NC, HW and SIL included zero (Fig. S10; Table S8).

Multivariate morphometric variation

PCA adequacy was high (KMO = 0.952; Bartlett $\chi^2(91) = 4326.3$, $P < 0.001$). PC1 explained 69.3% of variance and loaded positively across principal body-size measurements. It was the only component with an eigenvalue > 1 and was therefore the sole clustering input. Silhouette width was largest at the upper boundary of the tested range, k = 6 (0.542), followed by k = 5 (0.527) and k = 4 (0.527); k = 2 yielded 0.514. Bootstrap ARI for k = 6 was 0.723 ± 0.125. Because clustering was based on a single general-size coordinate and the optimum occurred at the upper boundary, these results reflect internally reproducible subdivisions of a continuous size gradient rather than evidence for discrete morphotypes. The gradient increased strongly with age and differed by sex (Fig. S8).

Live-weight associations and prediction

All 14 non-horn measurements correlated positively with live weight (all $P < 0.001$; Fig. 2; Table S9). NC was the leading pooled correlate (r = 0.800, 95% CI 0.754–0.838), followed by SIL (r = 0.795) and HB (r = 0.772); CG ranked fourth (r = 0.770). The pooled NC-only equation was LW = -13.729 + 1.999×NC, with the negative intercept rendered algebraically in Table S9.

On identical repeated partitions, NC alone yielded $R^2 = 0.633$, RMSE = 10.08 kg, and MAE = 8.19 kg, compared with $R^2 = 0.586$, RMSE = 10.70 kg, and MAE = 8.57 kg for CG (Table 3). However, the paired NC-minus-CG RMSE difference was − 0.62 kg (95% CI -1.70 to 0.56), and the MAE interval also included zero. CG had a numerically higher R^2 in males and females, whereas NC had a higher R^2 in the pooled older-than-12-month subset and among older females. No within-sex paired interval excluded zero (Table S17; Fig. S9). Thus, NC was the leading pooled predictor but not a consistently superior replacement for CG.

The field-oriented three-measurement equation (Table 3) was:

LW (kg) = − 56.518 + 0.996×NC + 0.669×SIL + 0.312×CG,

where NC, SIL and CG are in centimetres. It retained most of the fixed six-measurement model's repeated-CV performance (R^2 0.780 versus 0.794; RMSE 7.79 versus 7.54 kg). Its development ranges were NC 16.4–53.4 cm, SIL 42.8–95.7 cm and CG 49.6–129.1 cm.

Table 3
Fixed live-weight equations and repeated 10 × 10-fold cross-validated performance

Model	Equation	Applicable ranges	CV R^2	RMSE	MAE	Median APE (%)
NC only	LW = − 13.729 + 1.999×NC	NC 16.4–53.4 cm	0.633	10.08	8.19	14.5
NC + SIL	LW = − 50.563 + 1.221×NC + 0.855×SIL	NC 16.4–53.4; SIL 42.8–95.7 cm	0.762	8.11	6.37	11.2
NC + SIL + CG	LW = − 56.518 + 0.996×NC + 0.669×SIL + 0.312×CG	NC 16.4–53.4; SIL 42.8–95.7; CG 49.6–129.1 cm	0.780	7.79	6.21	10.8
Full stepwise	LW = − 64.087 + 0.980×NC + 0.494×SIL + 0.232×CG + 0.191×PBL + 1.409×CC − 0.330×NL	NC 16.4–53.4; SIL 42.8–95.7; CG 49.6–129.1; PBL 65.3–152.3; CC 7.3–15.4; NL 17.6–46.1 cm	0.794	7.54	5.95	9.7

LW is in kg, and all predictors are in cm. Ranges are observed minima–maxima. Equations should not be extrapolated. Repeated-fold means are shown; SDs and coefficient CIs appear in Supplementary File 1.

The overall six-measurement equation had an apparent R^2 = 0.807 (Table 4). With selection repeated inside the 10 outer folds, R^2 was 0.790, RMSE 7.63 kg, MAE 6.02 kg, MAPE 13.2% and median absolute percentage error 9.9% (Fig. 3). All six final predictors were selected in every outer fold; HL entered two folds. The negative NL coefficient is a conditional multivariable coefficient estimated from correlated predictors and should not be interpreted causally.

Table 4
Stepwise models with predictor selection repeated inside validation folds

Stratum	n	Selected full-sample predictors	Apparent R^2	Nested R^2	RMSE	MAE	Median APE (%)
Overall	287	NC, SIL, CG, PBL, CC, NL	0.807	0.790	7.63	6.02	9.9
Males	177	SIL, CG, NC, CC, NL	0.843	0.820	7.46	5.82	8.8
Females	110	CG, HL, NC, PBL	0.648	0.536	7.67	6.16	12.7

Predictor lists describe the model selected from each complete stratum; selection was repeated independently within every training fold. Age-stratum models appear in Supplementary File 1.

Calibration of held-out predictions yielded an observed-on-predicted intercept of 0.58 kg and a slope of 0.988. Mean predicted-minus-observed bias was 0.03 kg, with 95% limits of agreement − 14.95 to 15.01 kg. Bias varied with weight (trend $P < 0.001$), from + 3.44 kg in the lightest tertile to -4.69 kg in the heaviest. Leave-one-department-out pooled R^2 was 0.744 (Table S15).

Machine-learning benchmark and restricted-age sensitivity

Using the same 14 predictors and repeated folds, linear regression achieved R^2 = 0.799, RMSE = 7.45 kg and MAE = 5.94 kg. Random forest reduced MAE to 5.61 kg but slightly lowered R^2 to 0.794 and increased RMSE to 7.54 kg. Gradient boosting was similar (R^2 0.793, RMSE 7.56 kg), while the specified support-vector model performed less well (R^2 0.779). Therefore, no configuration improved all principal metrics (Table S19).

Excluding the 37 sheep aged ≤ 12 months, there were 250 sheep strictly older than 12 months (156 males, 94 females). Mean live weight was 54.59 ± 14.81 kg. The same six predictors were selected; the nested R^2 was 0.714, the RMSE was 7.90 kg, and the median absolute percentage error was 9.8%. PC1 explained 62.1%, and two components had eigenvalues > 1; in this subset, k = 2 had the largest silhouette (0.442). These shifts reinforce the sensitivity of cluster-number results to sample composition, while the broad size and prediction patterns persisted (Tables S20–S22).

Discussion

This study provides a combined morphometric and barymetric assessment of phenotypically identified Waila sheep sampled across four departments of northern Cameroon. It also corrects the analytical age boundaries, distinguishes protocol-specific PBL from conventional body/trunk length, and evaluates model performance using the audited individual-animal data. Direct size ranking against other tropical sheep populations remains inappropriate because age, sex, management and measurement protocols differ. Poulfouli sheep from the same broad agroecological zone provide useful regional context (Douandji et al. 2025), but represent a different named population.

Sex-associated differences were strongly age-dependent. At ≤ 12 months, adjusted contrasts were imprecise and did not show heavier males, whereas live-weight and NC differences were clear in older classes and largest after 36 months. This pattern may reflect biological sexual dimorphism, but also the production system: males are commonly retained for fattening and sale, while females are retained for breeding. Age class, locality, nutrition, management and selective retention are therefore intertwined. Sex-specific morphometric relationships and equations have also been reported elsewhere (Baffour-Awuah et al. 2000; Avalos-Castro et al. 2023). The present contrasts are conditional descriptions, not causal estimates of intrinsic sex effects.

Multivariate variation was dominated by one general-size axis. The complete-sample silhouette maximum occurred at k = 6, not k = 2 as previously reported, and at the upper boundary of the tested range. The older-than-12-month sensitivity analysis instead favoured k = 2 after retaining two components. These results do not justify naming discrete morphotypes. A one-dimensional size coordinate can be divided reproducibly without implying biological subpopulations; silhouette and bootstrap stability assess internal partition properties rather than genetic or functional distinctness

(Rousseeuw 1987; Hennig 2007). Demonstrating population structure would require size-adjusted shape analyses, targeted sampling and preferably molecular data.

NC was the strongest pooled correlate and had a modest pooled RMSE advantage over CG, but the paired interval included zero, and the direction did not persist consistently within sex. In males and females, CG produced numerically higher R^2; among older females, NC did. The adjusted model found positive contributions of both measurements and no clear sex interaction for either slope. This resolves the main interpretive concern: NC is a useful candidate measurement in this dataset, but the evidence does not support replacing chest girth universally. Many sheep studies have identified chest or heart girth as the leading predictor (Thys and Hardouin 1991; Markos et al. 2023; Pelotshweu et al. 2025). Hypotheses about cervical musculature, adiposity or rumen-fill robustness remain untested because body condition, feeding state, fleece and tape tension were not standardised.

Repeating selection within each training fold limited selection-related optimism: overall R^2 declined from 0.807 in-sample to 0.790 out-of-sample. The three-measurement model retained most of the fixed six-measurement model's performance and is easier to apply, but its RMSE remained 7.79 kg. The tested tree and support-vector configurations did not improve all error criteria. Random forest reduced MAE modestly while worsening R^2 and RMSE, illustrating why model claims should not rest on a single metric (Megersa et al. 2025). These comparisons are configuration-specific; they do not establish a general limit on machine learning.

A median absolute percentage error near 10% means half of the held-out predictions had smaller relative errors, not that every prediction was accurate. For a 50-kg sheep, this is about 5 kg at the median, while the Bland–Altman limits approached ± 15 kg, and NC-only individual prediction intervals were about 40 kg wide. The equations, therefore, do not, by themselves, establish fitness for drug dosing, precise sale pricing, or decisions near a fixed weight threshold. They are approximate estimation tools pending that can be of great help in smallholder farms.

Limitations include convenience sampling during one seasonal window, lack of flock identifiers, phenotypic rather than molecular breed assignment, age estimation from owner report and dentition, unstandardised feeding and body condition, and absence of replicate measurements. Random animal-level folds may be optimistic for application to entirely new flocks when management or ancestry is shared (Roberts et al. 2017). Validation was internal and leave-one-department-out rather than prospective in an independent population.

In conclusion, Waila sheep morphometry in this sample was dominated by an age- and sex-associated size gradient. Male–female differences in live weight and principal measurements emerged mainly after 12 months. NC was the leading pooled single predictor but had no consistent within-sex superiority over CG. The three- and six-measurement equations provided useful approximate estimates, and no tested machine-learning configuration improved all principal metrics. These equations should be treated as candidate Waila-specific tools until validated in independent flocks, seasons, and management conditions.

Declarations

Acknowledgements

The authors thank the regional and departmental delegations of MINEPIA, the heads of the Zootechnical and Veterinary Centres, and participating sheep owners for their assistance.

Funding

The authors declare that no funds, grants or other support were received for conducting this study.

Competing interests

The authors have no relevant financial or non-financial interests to disclose.

Author contributions

Conceptualisation and supervision: Blaise Arnaud Hako Touko and Fernand Tendonkeng. Data collection: Temfack Mekenjio Franck and Nsiyapnze Katte Yato Katte. Statistical analysis: Nsiyapnze Katte Yato Katte, Temfack Mekenjio Franck and Shey Ndogmi Yoniwo. Writing – original draft: Nsiyapnze Katte Yato Katte, Shey Ndogmi Yoniwo and Temfack Mekenjio Franck. Writing – review and editing: all authors. All authors approved the manuscript.

Ethics approval

The study protocol was approved by the Scientific and Ethical Committee of the University of Dschang, and field activities were authorised by the relevant MINEPIA regional and departmental delegations. The research was conducted in accordance with Chapter 7.8 (Use of animals in research and education) of the WOAH Terrestrial Animal Health Code and the CIOMS–ICLAS International Guiding Principles for Biomedical Research Involving Animals (2012). All procedures were non-invasive.

Consent to participate

The owners of all animals gave informed consent before their animals were measured.

Data availability

All supporting material is provided in Supplementary Files.

References

1. Avalos-Castro R, Segura-Correa JC, Palacios-Espinoza A, Angulo C, Yam-Tze JB, Nulutagua-Hernandez MA, Romero-Santillan F (2023) Prediction of live weight by morphometric measurements in females and males of Criollo sheep of the Oaxacan Mixteca, Mexico. Trop Subtrop Agroecosyst 26. https://doi.org/10.56369/tsaes.4771
2. Baffour-Awuah O, Ampofo E, Dodoo R (2000) Predicting the liveweight of sheep by using linear body measurements. Ghana J Agric Sci 33:207–212. https://doi.org/10.4314/gjas.v33i2.1872
3. Bland JM, Altman DG (1986) Statistical methods for assessing agreement between two methods of clinical measurement. Lancet 327:307–310. https://doi.org/10.1016/S0140-6736(86)90837-8
4. Cawley GC, Talbot NLC (2010) On over-fitting in model selection and subsequent selection bias in performance evaluation. J Mach Learn Res 11:2079–2107
5. Cocquyt G, Driessen B, Simoens P (2005) Variability in the eruption of the permanent incisor teeth in sheep. Vet Rec 157:619–623. https://doi.org/10.1136/vr.157.20.619
6. Douandji TPL, Meutchieye F, Djoufack TY, Munpen MJ, Nsangou AS (2025) Phenotypic features of Poulfouli sheep breed in the Sudano-Sahelian zone of Cameroon. Sci Afr 30:e03016. https://doi.org/10.1016/j.sciaf.2025.e03016
7. FAO (2012) Phenotypic characterization of animal genetic resources. FAO Animal Production and Health Guidelines No. 11. Food and Agriculture Organization of the United Nations, Rome
8. FAO (2015) The second report on the state of the world's animal genetic resources for food and agriculture. Food and Agriculture Organization of the United Nations, Rome
9. Hamito D (2009) Estimation of weight and age of sheep and goats. Ethiopian Sheep and Goat Productivity Improvement Program, Technical Bulletin No. 23
10. Hennig C (2007) Cluster-wise assessment of cluster stability. Comput Stat Data Anal 52:258–271. https://doi.org/10.1016/j.csda.2006.11.025
11. Markos D, Masho W, Baye M, Bayou E (2023) Morphometric traits and stepwise regression of indigenous sheep in the southwestern region, Ethiopia. J Agric Food Res 14:100666. https://doi.org/10.1016/j.jafr.2023.100666
12. Megersa AG, Negash F, Hailu A, Melak A, Assefa A, Aseged T, Sinkie S (2025) Prediction of body weight in indigenous sheep using random forest regression, support vector regression, and classification and regression trees algorithms. Vet Anim Sci 30:100517. https://doi.org/10.1016/j.vas.2025.100517
13. Pelotshweu K, Bolowe MA, Mosweu NP, Thutwa K, Kgwatalala PM (2025) Prediction of live body weight from linear body measurements in sheep breeds of Botswana. Biodiversitas 26:2696–2705. https://doi.org/10.13057/biodiv/d260614
14. Roberts DR, Bahn V, Ciuti S et al (2017) Cross-validation strategies for data with temporal, spatial, hierarchical, or phylogenetic structure. Ecography 40:913–929. https://doi.org/10.1111/ecog.02881

15. Rousseeuw PJ (1987) Silhouettes: a graphical aid to the interpretation and validation of cluster analysis. J Comput Appl Math 20:53–65. https://doi.org/10.1016/0377-0427(87)90125-7
16. Tchouamo IR, Tchoumboué J, Thibault L (2005) Caractéristiques socio-économiques et techniques de l'élevage ovin dans les hauts plateaux de l'Ouest du Cameroun. Tropicultura 23:201–211
17. Thys E, Hardouin J (1991) Prediction of sheep body weight in markets in the Far North Cameroon. Livest Res Rural Dev 3(1), Article 1. https://www.lrrd.org/lrrd3/1/hardouin.htm
18. Varma S, Simon R (2006) Bias in error estimation when using cross-validation for model selection. BMC Bioinformatics 7:91. https://doi.org/10.1186/1471-2105-7-91
19. Wilson RT (1991) Small ruminant production and the small ruminant genetic resource in tropical Africa. FAO Animal Production and Health Paper No. 88. Food and Agriculture Organization of the United Nations, Rome

Figures

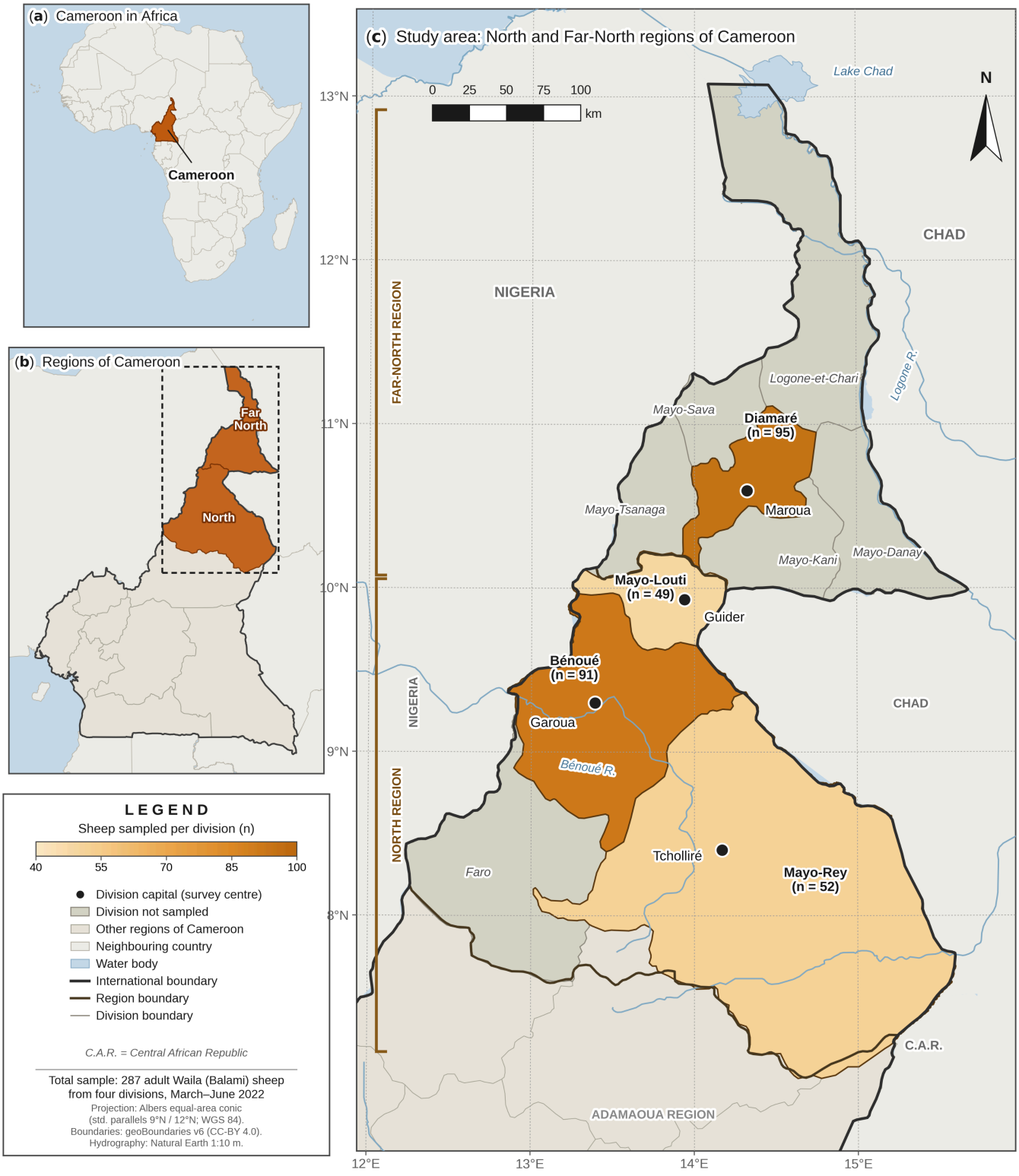


Figure 1

Departmental reference locations for sampling in the North and Far-North regions of Cameroon. Departmental capitals are geographic reference points and do not represent the locations of all participating flocks.

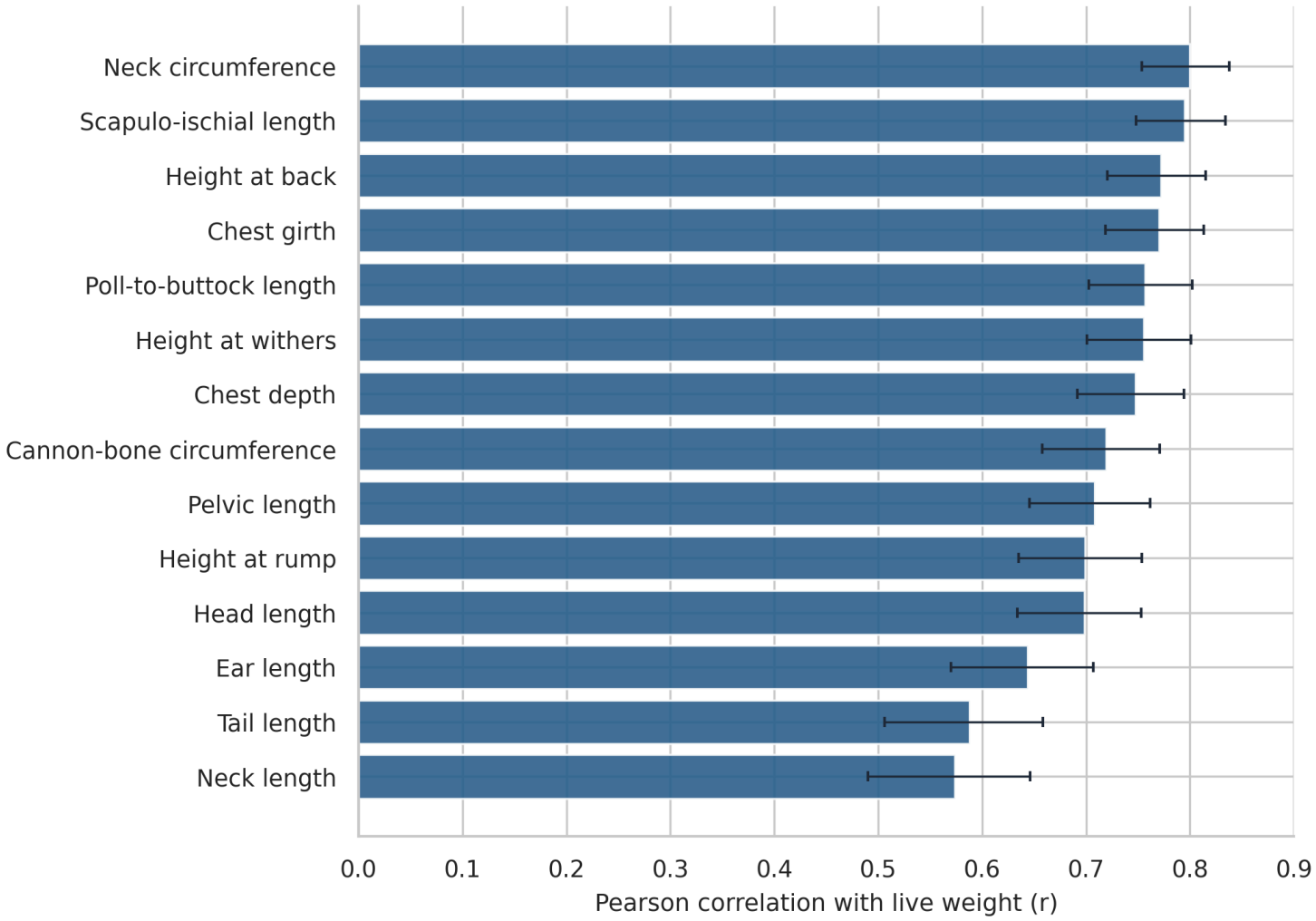


Figure 2

Pooled Pearson correlations and 95% confidence intervals between live weight and 14 non-horn measurements (n = 287). PBL is poll-to-buttock length, distinct from conventional trunk/body length.

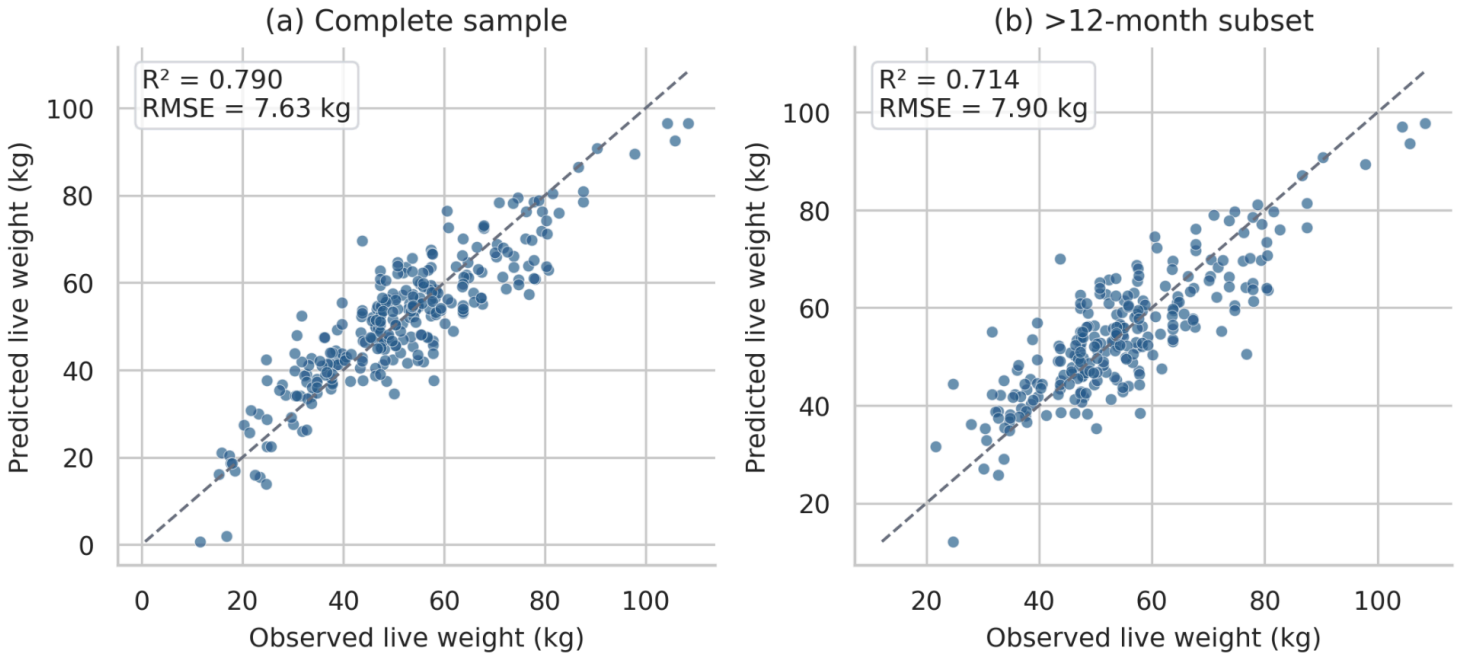


Figure 3

Held-out predictions from nested cross-validation with stepwise selection repeated in every training fold: (a) complete sample and (b) older-than-12-month subset. Dashed lines indicate equality.

Supplementary Files

This is a list of supplementary files associated with this preprint. Click to download.

- KatteetalWailaFinalSupplementaryMaterialTAHP.docx