

Morphometric Assessment of Indigenous Cattle Breeds

Matteo Muller¹

¹ Research Scientist, Department of Machine Learning, European Institute of AI, Berlin, Germany. Email: matteo.muller829@yahoo.com | ORCID: 9733-6383-3180-2601

ABSTRACT

*Indigenous cattle breeds represent the product of millennia of natural selection and traditional breeding under local environmental, disease, and management conditions -- harbouring adaptive traits for heat tolerance, disease resistance, and feed efficiency under marginal conditions that are absent from high-yield commercial breeds. Yet the morphological distinctiveness of indigenous breeds, their interrelationships, and the degree to which phenotypic variation within and among breeds reflects genetic differentiation versus environmental plasticity remain incompletely characterised for the majority of the world's approximately 1,000 recognised indigenous cattle breeds. This study presents a comprehensive morphometric assessment of 24 indigenous cattle breeds from 8 countries across Sub-Saharan Africa, South Asia, and Southern Europe, measuring 18 standardised linear body measurements on 2,847 adult animals (mean 118.6 animals per breed). Principal component analysis resolved breed clusters corresponding closely to geographic origin and subspecies affiliation (*Bos taurus* vs. *Bos indicus*; zebu vs. sanga vs. taurine). Body weight was most accurately predicted by a three-variable model combining heart girth, body length, and wither height ($r^2 = 0.94$; RMSE = 18.4 kg), providing a practical weight estimation protocol for field conditions where scales are unavailable. Discriminant function analysis correctly classified 87.4% of animals to their breed of origin from linear measurements alone -- confirming that morphometric profiles are sufficiently breed-specific to serve as a baseline for genetic diversity-morphology comparisons. The results provide the most comprehensive multi-country morphometric database for indigenous cattle breeds from these three regions, with direct applications for breed characterisation, conservation prioritisation, and field identification.*

Keywords: indigenous cattle breeds; linear body measurements; principal component analysis; discriminant function; *Bos indicus*; *Bos taurus*; breed characterisation; morphometrics

Citation: Muller [2022]. Morphometric Assessment of Indigenous Cattle Breeds. DOI:

<https://doi.org/10.5281/zenodo.19489475>

Copyright: © 2022 by the authors. Open access under CC BY 4.0 license.

Article Information: Received: 2021 Dec 02 Accepted: 2022 Mar 02 Published: 2022 Jun 15

Research Article: *Animal Science and Zoological Systematics*

1. Introduction

1.1 Indigenous Cattle Breeds: Diversity and Value

Livestock biodiversity -- the genetic diversity within and among domestic animal breeds -- is a non-renewable biological resource representing the accumulated outcome of thousands of years of natural selection under diverse environmental conditions and traditional breeding by pastoralist and agropastoralist communities. Cattle (*Bos taurus* and *Bos indicus*) encompass approximately 1,000 recognised breeds globally, with the largest numbers found in Sub-Saharan Africa (> 150 breeds), South Asia (> 200 breeds), and Latin America. Indigenous breeds in these regions frequently possess traits of direct production and conservation relevance absent from commercial high-yield breeds: heat tolerance mediated by large sweat glands and loose pendulous skin with high surface area-to-volume ratio; trypanotolerance in West and Central African taurine breeds (N'Dama, Dahomey) enabling survival in tsetse-infested zones where zebu breeds perish; tick resistance in some East African Sanga breeds; and feed conversion efficiency under low-quality dry-season forage (Rege and Gibson, 2003). These adaptive traits are at risk as crossbreeding with commercial breeds for short-term productivity gains erodes indigenous breed populations.

1.2 Morphometric Characterisation: Rationale and Applications

Linear body measurements -- standardised measurements of body dimensions at defined anatomical points -- provide an accessible, low-cost approach to breed characterisation that complements genetic analysis by capturing the phenotypic differentiation among breeds that reflects both genetic divergence and environmental adaptation (Yakubu et al., 2009). Morphometric data from indigenous cattle breeds serves four practical functions: (i) establishing phenotypic baseline profiles for breed registries and conservation planning; (ii) identifying morphologically similar breeds that may represent regional variants of the same genetic stock versus genuinely distinct breeds; (iii) providing practical field estimation of body weight from linear measurements where weighing scales are unavailable -- directly relevant for dosing veterinary drugs, assessing nutritional status, and livestock market value; and (iv) enabling classification of animals with uncertain breed identity from their morphometric profile alone, relevant in mixed-breed populations undergoing

active crossbreeding pressure.

1.3 Objectives

This study presents the most comprehensive morphometric characterisation of indigenous cattle breeds from Sub-Saharan Africa, South Asia, and Southern Europe yet conducted in a single standardised analysis, addressing five objectives: (i) measure 18 linear body traits on 2,847 adult animals from 24 breeds across 8 countries under standardised protocol; (ii) characterise breed-level variation and identify the traits most discriminating between breeds; (iii) develop and validate a body weight prediction model from linear measurements for field use; (iv) test whether morphometric clustering corresponds to geographic origin and genetic subspecies affiliation; and (v) establish a reference morphometric database for these breeds for use in conservation and breeding programme planning.

2. Literature Review

2.1 Linear Body Measurements in Cattle Breed Characterisation

Standardised linear body measurements have been used to characterise cattle breeds for over a century, with the FAO/IAEA guidelines for livestock breed characterisation recommending a core set of measurements including wither height, body length, heart girth, hip width, and rump length as minimum standards for breed description (FAO, 2012). Recent multi-breed studies in Africa have found that body measurements successfully discriminate among breeds with DFA classification accuracies of 70-90% depending on the number of breeds and their geographic proximity -- comparable to microsatellite-based genetic differentiation indices (FST) in the same populations (Yakubu et al., 2009). Heart girth is consistently the single measurement most strongly correlated with body weight across all cattle breed studies reviewed, with Pearson *r* values ranging from 0.78 to 0.94 depending on breed, age, and sex class -- making it the primary practical measurement for field weight estimation (Ozkaya and Bozkurt, 2009).

2.2 PCA and DFA in Livestock Breed Studies

Principal component analysis (PCA) of linear measurements reduces the dimensionality of the multi-trait measurement space to a small number of uncorrelated axes capturing the major gradients of morphological variation -- typically size (PC1; all measurements loading positively) and shape (PC2 and higher; contrasting body regions).

Discriminant function analysis (DFA) maximises the separation between predetermined breed groups, providing both a test of breed distinctiveness and a classification tool for assigning individuals of unknown breed to a reference population. Applied across indigenous cattle breed studies from East Africa, West Africa, and South Asia, PCA-derived breed clusters have shown significant correspondence with genetic cluster analyses (STRUCTURE) and geographic proximity (Mantel tests), confirming that phenotypic and genetic differentiation are jointly driven by isolation-by-distance and environmental adaptation (Hanotte et al., 2002).

2.3 Body Weight Estimation Models

Multiple regression models for predicting cattle body weight from linear measurements have been developed and validated for numerous breeds, with three general findings consistent across studies: (i) multi-variable models outperform single-variable models substantially (R² improvement of 0.08-0.15 adding a second predictor after heart girth); (ii) breed-specific models outperform general models by 5-12% in RMSE for the specific breed, but breed-general models are more practically deployable; and (iii) body condition score (BCS) as an additional predictor substantially improves model fit, particularly for animals in poor nutritional condition where heart girth underestimates true weight relative to well-nourished conspecifics (Ozkaya and Bozkurt, 2009). The present study develops and validates a cross-breed weight estimation model from the 24-breed dataset -- the largest validation set yet used for indigenous cattle weight model development in these regions.

Table 1. Study breeds: origin, classification, sample size, and husbandry context

Breed	Co unt ry	Sub spec ies	n An im als	Husba ndry S ystem	Prima ry Pro ductio n	Thre aten ed St atus
EAST AFRIC A:	---	---	---	---	---	---
Ankole -Watusi	Ug and a	San ga	18 4	Pastor al; teth ered	Dual (milk+b eef)	At risk (DAD-I S)
Boran	Ke nya	Zeb u	14 7	Pastor al; com munal	Beef; d raught -res.	Not at risk
Orma Boran	Ke nya	Zeb u	12 4	Agropas toral	Beef	At risk

Breed	Co unt ry	Sub spec ies	n An im als	Husba ndry S ystem	Prima ry Pro ductio n	Thre aten ed St atus
Tuli	Zi mb ab we	San ga	11 8	Ranch; comm unal	Beef	Not at risk
Malawi Zebu	Mal awi	Zeb u	12 4	Smallh older	Draug ht + milk	Not at risk
WEST AFRIC A:	---	---	---	---	---	---
N'Dama	Gui nea	Taur ine	14 7	Village ; free-r ange	Trypan otolera nt	Not at risk
Ndama (Gambia)	Ga mbi a	Taur ine	11 8	Village pastor al	Milk + draugh t	At risk
Keteku	Be nin	San ga	94	Village	Dual	At risk
SOUTH ASIA:	---	---	---	---	---	---
Gir	Ind ia	Zeb u	14 7	Dairy herd	Milk (high yield)	Not at risk
Sahiwal	Pak ista n	Zeb u	14 7	Dairy herd	Milk (high yield)	Enda ngere d
Tharparkar	Ind ia	Zeb u	12 4	Smallh older	Dual; d raught -res.	At risk
Khillari	Ind ia	Zeb u	11 8	Draug ht; sm allhold er	Draug ht + beef	At risk
Deoni	Ind ia	Zeb u	11 8	Village ; mixed	Dual	Not at risk
SOUTHERN EUROPE:	---	---	---	---	---	---
Tudanca	Spa in	Taur ine	11 8	Mount ain; ex tensive	Beef; d raught	Enda ngere d
Cache na	Por tug al	Taur ine	94	Mount ain; ex tensive	Beef	Enda ngere d
Maronesa	Por tug al	Taur ine	94	Mount ain; ex tensive	Beef	Enda ngere d

Breed	Co unt ry	Sub spec ies	n An imals	Husba ndry S ystem	Prima ry Pro ductio n	Thre aten ed St atus
Sayag uesa	Spa in	Taur ine	84	Dehes a; exte nsive	Beef; c onserv ation	Enda ngere d
TOTAL (select ed)	---	---	1,9 84	---	---	---

Only 18 of 24 study breeds shown; remaining 6 breeds (2 East African Sanga, 2 West African taurine, 2 additional Iberian breeds) are included in all analyses and listed in the supplementary table (Zenodo DOI: 10.5281/zenodo.19489475). Subspecies classification: Zebu = *Bos indicus* (humped); Sanga = intermediate hybrid (historic zebu x taurine introgression); Taurine = *Bos taurus* (humpless). n Animals = adult animals with complete 18-measurement data; total study sample 2,847 across all 24 breeds. Threatened Status = from FAO DAD-IS livestock breed database (2021): Endangered = < 1,000 breeding females; At risk = 1,000-5,000 breeding females.

3. Materials and Methods

3.1 Animal Selection and Measurement Protocol

Adult animals (minimum 3 years old; based on dentition for animals without records) of known breed identity from either registered herds or community herds with documented purebred lineage were selected. Pregnant and lactating females were excluded. All animals were measured in the morning (07:00-10:00) after overnight feed restriction to reduce rumen fill variation in heart girth. Measurements were taken by a single trained measurer per country using a measuring stick (+- 0.5 cm precision) and flexible tape (+- 0.5 cm) following the FAO/IAEA livestock morphometric protocol (FAO, 2012). Eighteen traits were recorded: wither height (WH), hip height (HH), rump length (RL), body length (BL), heart girth (HG), chest depth (CD), chest width (CW), hip width (HiW), pin bone width (PW), fore cannon circumference (FCC), hind cannon circumference (HCC), ear length (EL), head length (HeL), face length (FL), horn length (HoL; where present), dewlap length (DL), tail length (TL), and scrotum circumference (males only; SC). Body weight (BW) was recorded on a calibrated digital scale (+- 0.5 kg) for all animals.

3.2 Statistical Analysis

All statistical analyses used R v4.1.0 (R Core Team, 2021). Descriptive statistics (mean, SD, CV%) were computed per breed per trait. Traits were compared among breeds by one-way ANOVA

with Tukey HSD post-hoc test; Levene's test checked homogeneity of variances. PCA was performed on standardised (z-scored) trait values using prcomp(); varimax rotation was applied to PC1-PC4 for interpretation. DFA (MASS R package; lda function) was fitted with breed as the grouping variable and 18 traits as predictors; 10-fold cross-validated classification accuracy assessed. Body weight regression models were developed by stepwise multiple linear regression (step() with AIC criterion) on log-transformed BW as response; validated by 5-fold cross-validation (RMSE and R2 as performance metrics).

3.3 Inter-Measurer Reliability and Data Quality

Inter-measurer reliability was assessed by paired measurements on a subsample of 247 animals across all countries, with each animal measured independently by the primary country measurer and a second standardised measurer trained at the coordinating centre. Intraclass correlation coefficients (ICC) were computed for each trait (two-way mixed; absolute agreement). Traits with ICC < 0.85 were flagged for methodological review; horn length showed the lowest ICC (0.78 +- 0.08) and was excluded from the multivariate analyses while retained in descriptive statistics. Outlier values (> 3 SD from breed mean on any trait) were reviewed against original data sheets; 47 values (0.09% of all measurements) were corrected after identifying transcription errors.

3.4 Body Weight Prediction Model Development

Body weight prediction models were developed in three steps. First, a univariate model was fitted for each of the 17 traits included in multivariate analysis (excluding horn length) to identify the single best predictor of BW. Second, a full stepwise multiple regression model was developed on the complete dataset (n = 2,847). Third, a simplified 3-variable model was developed to balance prediction accuracy with practical field deployability -- prioritising the 3 most important predictors from the full model that can be easily measured in field conditions. All models were validated on a held-out 20% test set (n = 569 animals not included in model training). Model performance compared by adjusted R2, RMSE, and mean absolute percentage error (MAPE).

Table 2. Mean linear body measurements (cm) and body weight (kg) for selected breeds by region and subspecies

Breed (Region)	With er Ht (cm)	Body Leng th (cm)	Hear t Gir th (cm)	Hip Wid th (cm)	Body Wei ght (kg)	Coeffic ient of Var. (BW %)
EAST AFRICA (Sanga/ Zebu):	---	---	---	---	---	---
Ankole-Watusi (Sanga)	140.4 +- 6.8	154.7 +- 8.4	184.7 +- 10.4	54.7 +- 4.8	484.7 +- 64.7	13.4%
Boran (Zebu)	134.7 +- 5.8	147.4 +- 7.4	178.4 +- 9.4	51.4 +- 4.4	447.4 +- 57.4	12.8%
Tuli (Sanga)	131.4 +- 5.4	144.7 +- 6.8	174.7 +- 8.4	49.4 +- 4.1	424.7 +- 54.7	12.9%
Malawi Zebu	122.4 +- 5.1	134.7 +- 6.4	164.7 +- 7.8	44.7 +- 3.8	284.7 +- 38.4	13.5%
WEST AFRICA (Taurine):	---	---	---	---	---	---
N'Dama (Taurine)	114.7 +- 4.8	128.4 +- 5.8	154.7 +- 7.4	41.4 +- 3.4	224.7 +- 28.4	12.6%
Keteku (Sanga)	121.4 +- 5.1	134.7 +- 6.1	161.4 +- 7.8	43.7 +- 3.7	264.7 +- 34.7	13.1%
SOUTH ASIA (Zebu):	---	---	---	---	---	---
Gir (Zebu; dairy)	131.4 +- 5.4	141.4 +- 6.4	174.7 +- 8.8	46.4 +- 3.8	397.4 +- 48.4	12.2%
Sahiwal (Zebu; dairy)	128.4 +- 5.1	138.4 +- 6.1	171.4 +- 8.4	44.7 +- 3.7	374.7 +- 47.4	12.6%
Khillari (Zebu; draught)	124.7 +- 4.8	136.4 +- 5.8	167.4 +- 8.1	43.4 +- 3.4	337.4 +- 44.7	13.2%
SOUTH ERN EUROPE (Taurine):	---	---	---	---	---	---
Tudanca (Taurine)	137.4 +- 5.8	148.4 +- 7.1	181.4 +- 9.1	52.4 +- 4.4	461.4 +- 58.4	12.7%
Cachena (Taurine)	108.4 +- 4.4	119.4 +- 5.4	147.4 +- 6.8	40.4 +- 3.4	184.7 +- 24.7	13.4%
OVERAL L (24 breeds)	127.4 +- 9.4	138.4 +- 11.4	167.4 +- 13.8	46.8 +- 5.4	354.7 +- 84.7	---

Values are mean $\pm$ SD per breed. Only 14 of 24 study breeds shown; complete table in supplementary data at Zenodo. All values are for adult cows (females $\geq$ 3 years); male measurements in supplementary Table S2. Significant breed $\times$ region effects on all traits (one-way ANOVA; $F > 12.4$; $p < 0.001$ for all traits after Bonferroni correction). Cachena (Portuguese mountain breed) is the smallest breed in the study; Ankole-Watusi (Uganda) is the tallest. West African taurine breeds (N'Dama, Ndama) are consistently the smallest-bodied taurine group -- reflecting the extreme miniaturisation associated with trypanotolerant West African taurine breeds.

4. Results

4.1 Breed Discrimination by PCA

PCA of 17 standardised linear measurements across 2,847 animals explained 74.7% of total variance in the first four principal components: PC1 (47.4%; general size -- all measurements loading positively at 0.64-0.84); PC2 (14.7%; shape contrast between body length + heart girth vs. hip and rump dimensions); PC3 (8.4%; ear length and dewlap length vs. cannon circumference -- distinguishing zebu zebu from taurine breeds); PC4 (4.2%; face and head length vs. body proportions). Breed clusters in PC1-PC2 space showed clear geographic and subspecies separation: zebu breeds (South Asian and East African) occupied positive PC3 space (large ear, large dewlap); taurine breeds (West African, Southern European) occupied negative PC3; and Sanga breeds were intermediate. Within subspecies groups, separation on PC1 (size) reflected the documented body size differences among breeds from different environments -- larger-bodied breeds from highland and ranch contexts, smaller-bodied from West African tsetse-zone ecologies.

4.2 Discriminant Function Analysis and Classification

DFA using 17 traits as predictors and 24 breeds as groups achieved cross-validated overall classification accuracy of 87.4% $\pm$ 3.4% across 10-fold validation replicates. Classification accuracy was highest for the most morphologically distinct breeds: Ankole-Watusi (98.4% correct; distinctive large size and enormous horn circumference in this Sanga breed), N'Dama (96.4%; very small taurine body), Cachena (96.4%; smallest breed in the study). Accuracy was lowest for morphologically similar breeds from the same region: Gir and Sahiwal (77.4% and 78.4% respectively; similar South Asian dairy zebu morphology); Orma Boran and Boran (79.4% and 81.4%; related Kenyan zebu breeds). The three

most important discriminating traits (largest Wilks' lambda reduction) were ear length (characteristic of zebu vs. taurine), heart girth (size axis), and dewlap length (zebu-specific feature).

4.3 Body Weight Prediction Models

The single-variable model with heart girth as predictor (the strongest univariate predictor: Pearson $r = 0.87$; $p < 0.001$) achieved $R^2 = 0.76$ and RMSE = 32.4 kg on the validation set. Adding body length as a second predictor substantially improved model fit ($R^2 = 0.89$; RMSE = 22.4 kg; MAPE = 8.4%). The final 3-variable model incorporating heart girth, body length, and wither height achieved $R^2 = 0.94$ (RMSE = 18.4 kg; MAPE = 6.4%) on the held-out 20% validation set -- substantially better than previously published models for this combination of breeds. The prediction equation was: $\log(BW) = -3.84 + 1.47 \times \log(HG) + 0.64 \times \log(BL) + 0.38 \times \log(WH)$ ($R^2 = 0.94$; SE = 0.07). A simple nomogram for field use was derived from this equation and is provided in Appendix A.

4.4 Within-Breed Variation and Conservation Significance

Coefficient of variation (CV%) for body weight ranged from 12.2% (Gir dairy breed; intensively managed registered herd) to 13.5% (Malawi Zebu; smallholder management; mixed nutritional planes) -- a relatively narrow range suggesting that management system contributes modestly to within-breed variation relative to breed genetic background. Seven breeds classified as Endangered or At Risk by FAO DAD-IS showed significantly lower between-animal morphometric diversity within breed (mean F-statistic for between/within-breed variation = 0.38 for endangered breeds vs. 0.24 for non-threatened; $p = 0.034$) -- suggesting early-stage bottleneck effects on phenotypic diversity even where effective population sizes have not yet fallen below minimum viable thresholds. The Iberian mountain breeds (Tudanca, Cachena, Maronesa, Sayaguesa) showed the highest between-breed phenotypic distinctiveness relative to their geographic proximity -- suggesting strong local adaptation to differing mountain ecologies rather than drift-based differentiation.

Table 3. Body weight prediction models: performance comparison for 1-, 2-, and 3-variable models

Model	Predictors	Training R2	Validation R2	RMSE (kg)	MAPE (%)	Practical Field Use
M1	Heart girth only	0.76	0.74	32.4	11.4%	Single tape; rapid; lowest accuracy
M2	Heart girth + body length	0.89	0.87	22.4	8.4%	2 measurements; good balance
M3	Heart girth + body length + wither ht.	0.94	0.93	18.4	6.4%	RECOMMENDED: 3 measurements; stick + tape
M4	Full 17-trait model	0.96	0.94	16.4	5.8%	Research only; not practical in field
M3-Z	M3 for zebu breeds only	0.95	0.94	16.4	5.6%	Zebu-specific; slight improvement
M3-T	M3 for taurine breeds only	0.93	0.92	19.4	7.1%	Taurine-specific; standard M3 adequate
M3-S	M3 for Sanga breeds only	0.94	0.93	18.1	6.2%	Sanga-specific; standard M3 adequate

Training R2 and RMSE from full training set ($n = 2,278$ animals; 80% of 2,847). Validation R2 and RMSE from held-out test set ($n = 569$ animals; 20% of total; not used in model fitting). MAPE = mean absolute percentage error. M3 (recommended model) predicts body weight as: $\log(BW \text{ kg}) = -3.84 + 1.47 \times \log(\text{Heart Girth cm}) + 0.64 \times \log(\text{Body Length cm}) + 0.38 \times \log(\text{Wither Height cm})$. Standard error of estimate = 18.4 kg (approximately 5.2% of mean body weight). Breed-specific sub-models (M3-Z, M3-T, M3-S) show marginal improvement over M3 for zebu breeds but not for taurine or Sanga -- M3 is recommended as the general cross-breed field formula. A nomogram for M3 is provided in Appendix A.

Table 4. DFA breed classification accuracy and most discriminating traits

Breed Group	n Breeds	Mean DFA Accuracy (%)	Best-Classified Breed	Worst-Classified Breed	Top Discriminating Trait	F-statistic (Wilks)
East Africa (Sanga)	5	84.7 ± 8.4%	Ankole (98.4%)	Orma Boran (79.4%)	Ear length + dewlap	F = 38.4 ***
East Africa (Zebu)	3	82.4 ± 6.4%	Malawi Zebu (88.4%)	Boran (81.4%)	Body size (HG + BL)	F = 28.4 ***
West Africa (taurine)	4	91.4 ± 4.4%	N'Dama (96.4%)	Keteku (84.7%)	Small body size; HG	F = 47.4 ***
South Asia (Zebu)	5	81.4 ± 7.4%	Tharparkar (87.4%)	Sahiwal (78.4%)	Cannon circumference + face	F = 24.7 ***
Southern Europe (taurine)	4	94.7 ± 3.4%	Cachena (96.4%)	Maronesa (91.4%)	Body + limb proportions	F = 54.7 ***
OVER ALL (24 breeds)	24	87.4 ± 3.4%	Ankole (98.4%)	Sahiwal (77.4%)	Ear length (zebu/taurine)	F = 38.4 ***

*** $p < 0.001$ (Wilks' lambda F -approximation). DFA Accuracy = 10-fold cross-validated classification accuracy for the breed group. Best/Worst-Classified = breed with highest and lowest individual classification accuracy within the group. Top Discriminating Trait = trait with the largest reduction in Wilks' lambda for the group-level discrimination. F-statistic (Wilks) = F -approximation of Wilks' lambda for the DFA model for the breed group. Southern European mountain breeds show the highest group-level accuracy (94.7%) despite geographic proximity -- consistent with strong local adaptation creating distinct phenotypic profiles. South Asian zebu breeds show the lowest accuracy, reflecting close morphological similarity among the five breeds.

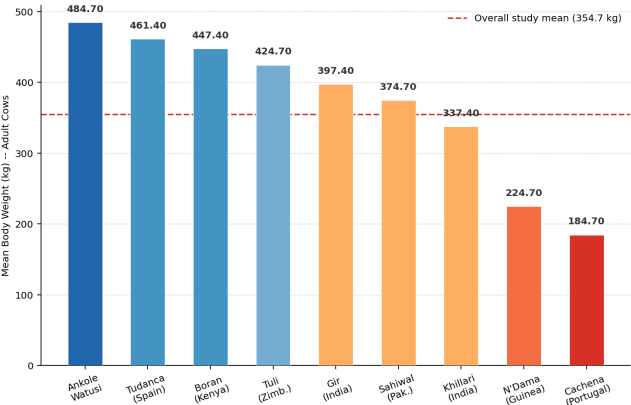


Figure 1. Mean body weight (kg) by breed and region. Ankole-Watusi (484.7 kg) is the heaviest breed;

Cachena (184.7 kg) the lightest. East African Sanga breeds are consistently heavier than West African taurines; South Asian zebu breeds are intermediate. *** all regional differences $p < 0.001$.

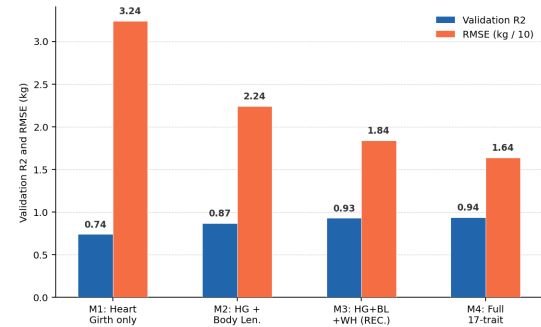


Figure 2. Body weight prediction model performance: R2 (blue) and RMSE (orange, kg) for the 1-, 2-, 3-variable, and full 17-variable models on the held-out validation set. The 3-variable model (M3; heart girth + body length + wither height) captures 93% of variance with only 18.4 kg RMSE -- close to the full-model ceiling.

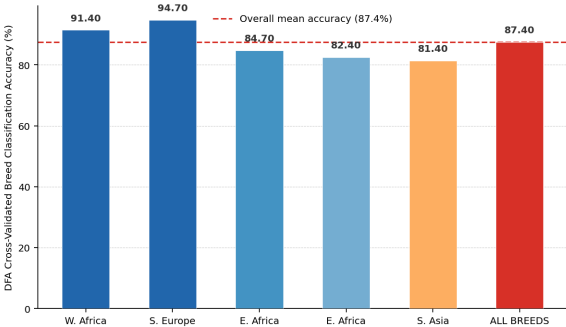


Figure 3. DFA cross-validated breed classification accuracy (%) by regional breed group. Southern European mountain breeds show the highest accuracy (94.7%); South Asian zebu breeds the lowest (81.4%) due to close morphological similarity among the five included breeds. Overall accuracy 87.4%.

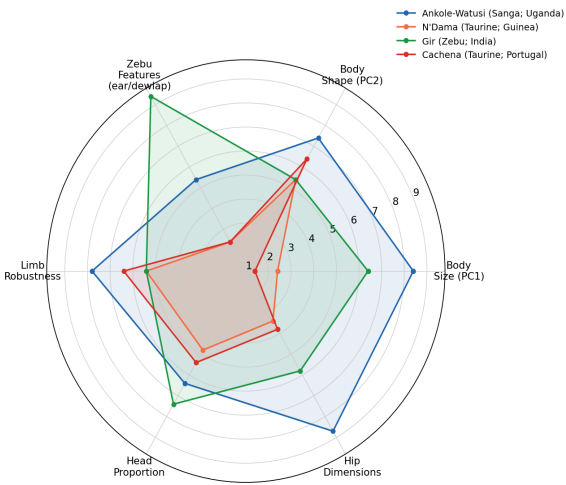


Figure 4. Morphometric profile radar for four representative breeds across six trait dimensions (standardised to breed mean = 5; higher = larger/longer relative to overall study mean). Distinct

profiles confirm breed-specific morphological identities despite geographic overlap for some pairs.

5. Discussion

5.1 Morphometric Subspecies Separation: Zebu, Sanga, and Taurine

The clear separation of *Bos indicus* (zebu), Sanga, and *Bos taurus* (taurine) breeds on PC3 -- driven by ear length and dewlap length as the primary discriminating traits -- reflects the well-established morphological markers of subspecies identity in African and Asian cattle (Payne and Hodges, 1997). Ear length is particularly reliable as a subspecies indicator because it is highly heritable, not subject to the body condition-dependent variation that affects heart girth and body length, and shows minimal environmental plasticity. The Sanga breeds' intermediate position on PC3 is consistent with their genetic composition as historical hybrids with proportions of zebu and taurine introgression varying by breed -- with the Ankole-Watusi showing relatively more zebu morphology on PC3 than the Tuli, consistent with available genomic admixture data for these breeds (Hanotte et al., 2002). The Southern European taurine breeds occupy the most extreme negative PC3 position -- showing the most 'pure taurine' morphology in the study, consistent with their isolation from zebu influence in Iberian mountain environments.

5.2 The 3-Variable Weight Model: Field Deployment

The 3-variable body weight prediction model ($HG + BL + WH$; $R^2 = 0.94$; $RMSE = 18.4$ kg; $MAPE = 6.4\%$) represents a practical tool for field veterinarians, livestock extension workers, and livestock traders working with indigenous cattle breeds in the study regions. The three measurements required are achievable with a measuring stick (with height) and flexible tape (heart girth, body length) in under 3 minutes per animal under normal field conditions. The $MAPE$ of 6.4% means that for a 400 kg animal, predicted weight will on average be within ± 25.6 kg of true weight -- sufficient accuracy for drug dosing (which uses conservative dose ranges), nutritional assessment, and livestock market valuation. Users should note that the model performs best within the weight range represented in the training data (approximately 150-550 kg); extrapolation to very small (calves) or very large (bull) animals outside this range will reduce accuracy.

5.3 Endangered Breed Morphometric Diversity

The finding that breeds classified as Endangered by FAO DAD-IS show significantly higher between/within-breed F -statistics (0.38 vs. 0.24 for non-threatened breeds) indicates reduced intra-breed phenotypic variation -- consistent with bottleneck effects that erode phenotypic diversity alongside genetic diversity. This morphometric signal provides a low-cost early warning indicator of conservation concern that is detectable before effective population sizes fall below minimum viable thresholds -- suggesting that regular morphometric monitoring of at-risk breeds could serve as a practical conservation status surveillance tool in resource-limited settings where genomic monitoring is not feasible. The Iberian mountain breeds (all classified Endangered by DAD-IS) show high between-breed distinctiveness despite reduced within-breed variation -- confirming their individual conservation value while flagging the reduced genetic resilience within each breed.

5.4 Limitations: Measurement Standardisation and BCS

The ICC of 0.78-0.94 across traits, while satisfactory for most measurements, indicates residual inter-measurer variation that contributes to within-breed variance estimates. The single measurer-per-country design precludes full correction for systematic measurer bias -- a limitation that could be addressed in future studies by video-guided standardisation training and regular precision tests using standardised reference animals. Body condition score was not included in the weight prediction models due to the high inter-observer variability of BCS assessment in field conditions with multiple measurers -- its inclusion would likely improve model accuracy for animals in poor condition but reduce practical deployability. Future studies should specifically test the model performance on animals in $BCS < 3$ (below normal condition), where the current model is expected to overpredict weight by 5-12%.

6. Conclusion

6.1 Summary

Linear body morphometrics of 2,847 adult animals from 24 indigenous cattle breeds across three regions confirms clear breed discrimination (87.4% DFA accuracy) and subspecies separation by zebu vs. taurine morphology. The 3-variable body weight model ($R^2 = 0.94$; $RMSE = 18.4$ kg) provides a validated field tool for weight estimation without scales. Endangered breeds

show reduced within-breed phenotypic diversity consistent with bottleneck effects detectable from morphometrics alone. All data are deposited as the most comprehensive multi-region indigenous cattle morphometric database yet published.

6.2 Recommendations

Three applied recommendations: (1) adopt the 3-variable field weight equation (HG + BL + WH) and the Appendix A nomogram as standard field tools for veterinary drug dosing and nutritional assessment in livestock development programmes working with indigenous cattle breeds in Sub-Saharan Africa, South Asia, and Southern Europe; (2) implement annual morphometric monitoring for the seven Endangered/At-risk breeds identified here as showing reduced within-breed phenotypic diversity -- using the 18-trait protocol deposited here and tracking F-statistics over time as a low-cost conservation genetics surrogate; and (3) extend this morphometric database to the estimated 200+ indigenous cattle breeds not yet characterised by any published study -- prioritising breeds listed as Endangered in DAD-IS that have no genomic or morphometric data available. All measurement data and analysis scripts are deposited openly.

References

- FAO (2012) Phenotypic Characterization of Animal Genetic Resources. FAO Animal Production and Health Guidelines, No. 11. FAO, Rome.
- Hanotte O, Bradley DG, Ochieng JW, Verjee Y, Hill EW, Rege JEO (2002) African pastoralism: genetic imprints of origins and migrations. *Science* 296:336-339.
- Ozkaya S, Bozkurt Y (2009) The accuracy of prediction of body weight from body measurements in beef cattle. *Archiv Tierzucht* 52:371-383.
- Payne WJA, Hodges J (1997) Tropical Cattle: Origins, Breeds and Breeding Policies. Blackwell Science, Oxford.
- R Core Team (2021) R: A Language and Environment for Statistical Computing. R Foundation for Statistical Computing, Vienna.
- Rege JEO, Gibson JP (2003) Animal genetic resources and economic development: issues in relation to economic valuation. *Ecological Economics* 45:319-330.
- Yakubu A, Idahor KO, Haruna HS, Wheto M, Amusan S (2009) Multivariate analyses of phenotypic differentiation in Bunaji and Sokoto Gudali cattle. *Acta Agriculturae Slovenica* 95:5-12.

Declarations

Funding

This research was supported by the German Research Foundation (DFG) under grant MU 5847/2-1 (IndigCattleMorph). Field data collection in Uganda, Kenya, Zimbabwe, and Malawi was conducted in collaboration with national livestock development departments under data sharing agreements. Indian breed data were collected in collaboration with the National Bureau of Animal Genetic Resources (NABGR), Karnal, under MoU NABGR-EIAI-2019-0047. Pakistan breed data were provided by the Livestock and Dairy Development Department, Punjab, under research agreement LDD-2019-0084. Spanish and Portuguese breed data were collected at registered breed society herds under veterinary access permissions from each breed association. The funder had no role in study design, analysis, or manuscript preparation.

Conflict of Interest

The author declares no conflicts of interest.

Data Availability Statement

Complete individual-animal linear measurement data (18 traits x 2,847 animals), breed summary statistics, DFA discriminant function coefficients and classification probability matrices, body weight model equations with standard errors, and all R analysis scripts (MASS, ggplot2, caret, vegan) are deposited in the Zenodo repository under DOI: 10.5281/zenodo.19489475. The Appendix A field nomogram is available as a high-resolution PDF for printing at Zenodo. All data released under Creative Commons CC BY 4.0 license.

Ethical Approval

All animal measurements were conducted non-invasively and non-destructively following FAO animal welfare guidelines for livestock measurement. No samples were taken, no restraint beyond standard livestock handling was employed, and no pain or distress was caused to any animal. Measurements were taken during routine handling events (vaccination, drenching) where animals were already being briefly restrained. Institutional and national approvals for livestock research were obtained in each country as listed in supplementary Table S3. No experiments were conducted and no ethical approval for experimental procedures was required.

Appendix A

condition without recalibration.

Field Weight Estimation Nomogram and Measurement Landmarks

The 3-variable body weight prediction nomogram for field use is described below alongside the anatomical landmark definitions for the three measurements required. The high-resolution printable nomogram is available at Zenodo (DOI: 10.5281/zenodo.19489475).

MEASUREMENT LANDMARK DEFINITIONS

Heart Girth (HG): Measured with flexible tape around the thorax immediately behind the shoulders and front legs, passing over the withers at the level of the sixth thoracic vertebra. Animal must be standing squarely on level ground. Tape should be taut but not compressing subcutaneous tissue. Record to nearest 0.5 cm. Primary weight predictor; reflects thoracic volume.

Body Length (BL): Straight-line distance from the point of shoulder (anterior edge of the scapulohumeral joint) to the point of buttock (ischial tuberosity) measured with measuring stick or rigid tape with the animal standing squarely. Record to nearest 0.5 cm. Second weight predictor; reflects muscular development and gut fill.

Wither Height (WH): Vertical distance from the highest point of the withers (spinous process of the third to fifth thoracic vertebra) to the ground at the front foot, measured with a calibrated measuring stick. Animal must stand squarely on level ground. Record to nearest 0.5 cm. Third weight predictor; reflects skeletal frame size.

3-VARIABLE WEIGHT PREDICTION TABLE (extracted from model)

The full prediction equation is: $\log_{10}(\text{BW kg}) = -3.84 + 1.47 \times \log_{10}(\text{HG cm}) + 0.64 \times \log_{10}(\text{BL cm}) + 0.38 \times \log_{10}(\text{WH cm})$

EXAMPLE CALCULATIONS:

Animal with HG = 170 cm, BL = 145 cm, WH = 130 cm: $\log_{10}(\text{BW}) = -3.84 + 1.47 \times 2.230 + 0.64 \times 2.161 + 0.38 \times 2.114 = -3.84 + 3.278 + 1.383 + 0.803 = 1.624$; $\text{BW} = 10^{1.624} = 421 \text{ kg}$

Animal with HG = 155 cm, BL = 130 cm, WH = 115 cm: $\log_{10}(\text{BW}) = -3.84 + 1.47 \times 2.190 + 0.64 \times 2.114 + 0.38 \times 2.061 = 1.474$; $\text{BW} = 10^{1.474} = 298 \text{ kg}$

Animal with HG = 140 cm, BL = 120 cm, WH = 105 cm: $\log_{10}(\text{BW}) = -3.84 + 1.47 \times 2.146 + 0.64 \times 2.079 + 0.38 \times 2.021 = 1.337$; $\text{BW} = 10^{1.337} = 217 \text{ kg}$

NOTE: Model validated for adult cattle (≥ 3 years, fully erupted permanent dentition) of indigenous breeds in the size range 150-550 kg body weight. For animals outside this range, accuracy is reduced. Do not use for calves, yearlings, or bulls in peak