Comparing Methodologies for Deriving Methane Phenotypes from Laser Methane Detector Data in Small Ruminants

Emna Rekik, Maria Wurzinger, Johann Sölkner, Tesfaye Getachew, Barbara Rischkowsky, Dawit Solomon and Aynalem Haile

Presented by Sarah-Joe Burn

Laser Methane Detector (LMD):
What Does It Measure?

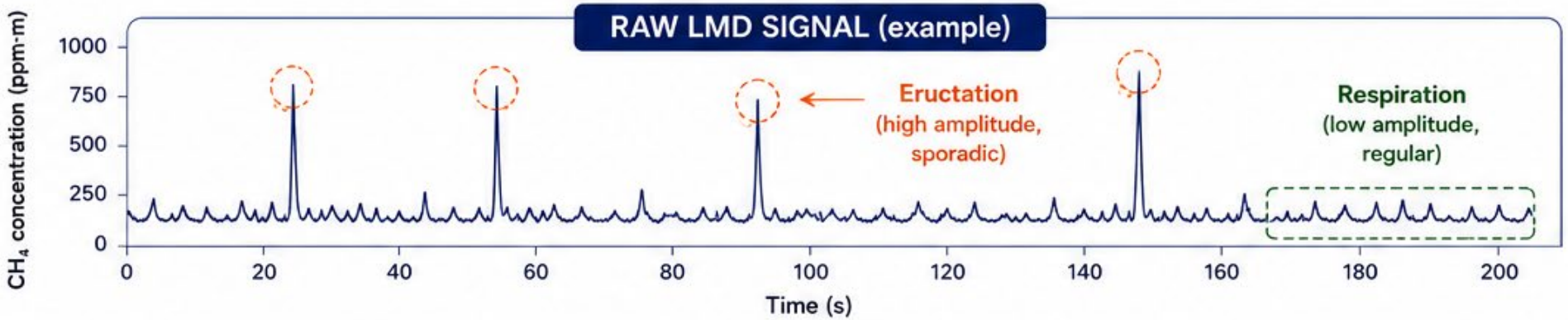
MEASUREMENT PRINCIPLE

- Infrared absorption spectroscopy
- Measures methane concentration along the laser path
- Output: CH_4 (ppm × m)

DEVICE CHARACTERISTICS

- Portable handheld device
- Non-contact measurement
- Distance: 0.5 – 30 m
- Range: 1 – 50,000 ppm × m

RAW LMD SIGNAL (example)



Eructation peaks
Large, intermittent methane releases

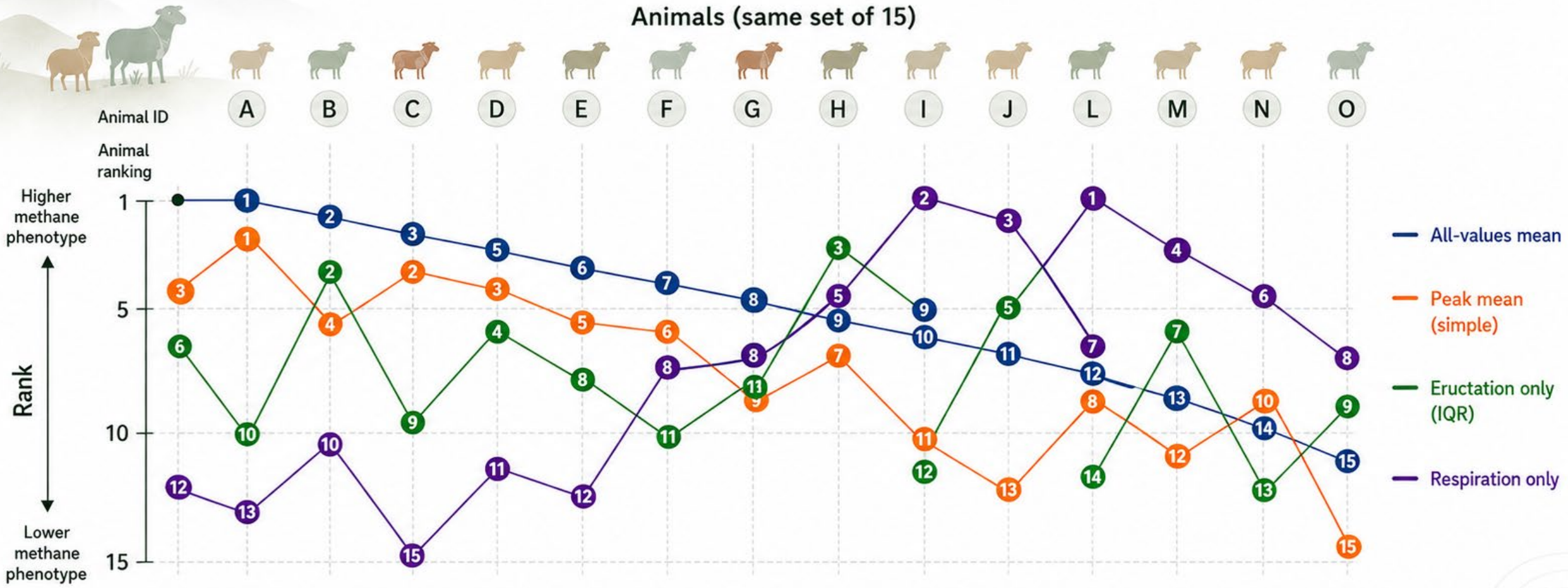
Respiration signal
Small, continuous oscillations

Many Methane Phenotypes from Raw LMD Data

One LMD recording can be processes in multiple ways to produce several methane phenotype

Same animal, Same Data : different ranking depending on the measured trait

Animal rank is not a biological fact — it is a methodological artifact.



Different analytical methodologies applied to the same raw LMD data can produce very different methane phenotypes and completely different animal rankings.

The objective of the study

To evaluate whether different published methodologies for deriving methane phenotypes from LMD raw data produce **equivalent**, **repeatable**, and **biologically relevant** phenotypes when applied to the same large-scale field dataset from sheep, and to assess the **stability of animal rankings** across methodologies.

We are not asking:

"Which methodology is best?"

We are asking:

"Which methodology best reflects biologically and genetically meaningful methane variation?"

How Do We Judge a "Good" Methane Phenotype?

No single validation criterion exists

Increasing relevance to breeding decisions

5 GENOMIC VALIDATION (FUTURE WORK)

- GWAS associations
- Genomic prediction accuracy
- Direct relevance to breeding outcomes

4 GENETIC PARAMETERS

- Heritability (h^2)

Additive genetic effect

More additive genetic signal relative to noise

3 SENSITIVITY TO BIOLOGICAL DRIVERS

- Diet composition
- Breed
- Body weight
- Lactation / feeding stage

Phenotype responds in the expected direction to known biological effects

2 AGREEMENT WITH REFERENCE METHODS

- Respiration chamber
- GreenFeed
- SF_6 (tracer)

Stronger agreement with external measures of methane (under their own limitations)

1 INTERNAL CONSISTENCY

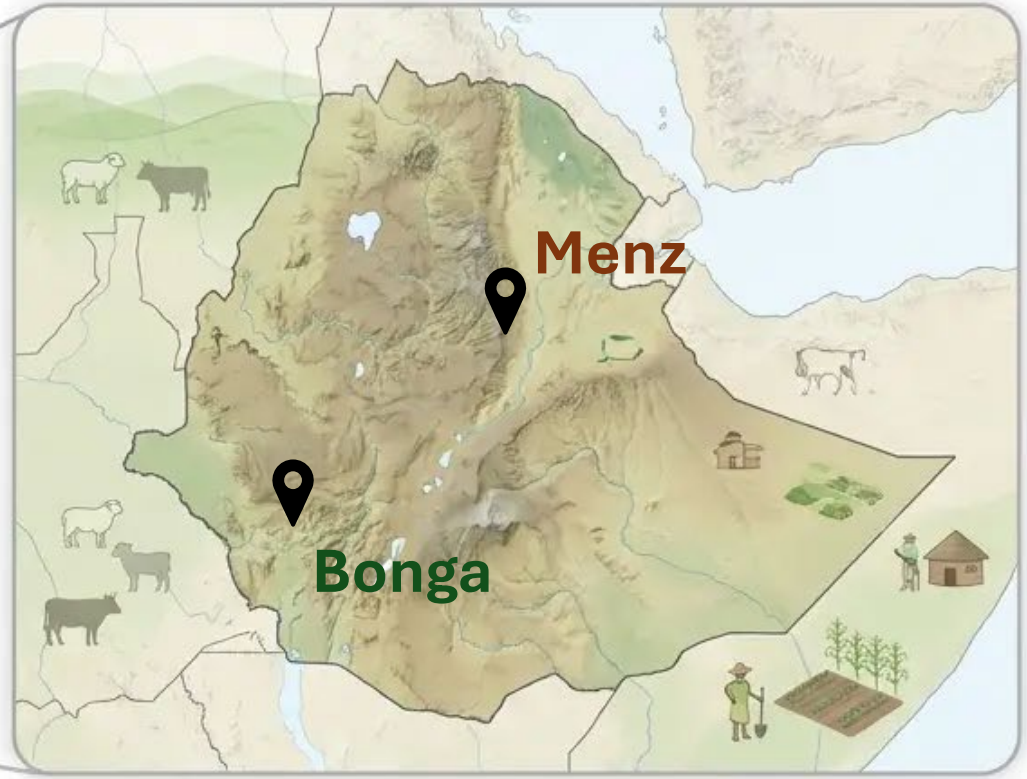
- Repeatability (within-method)
- Day-to-day reproducibility

Method is stable and reproducible, but not necessarily correct

IMPORTANT

High repeatability does not necessarily mean biological correctness.

MATERIALS AND METHODS



Ethiopia – East Africa

Over 80% of the population involved in smallholder agriculture

One of Africa's largest livestock populations

Diverse climate: cool highlands to semi-arid lowlands

Bonga breed

Bonga
303 (46.8%)

647 Animals from two indigenous Ethiopian breeds

Menz
344 (53.2%)

Menz breed

RECORDING PROTOCOL

- Each animal was measured **once a day**
- Each animal was measured for **three consecutive days**
- Recording duration ranged from **15 to 30 minutes** per animal
- Measurements were taken **1 meter** from the nostrils
- **Body weight** was recorded after methane measurements
- **Time since last feed access** was documented
- **lambs, ewes (0–11 parities), breeding rams, and castrated rams**
- Animals were **grazing** in the communal grazing area.

Field LMD measurement in Ethiopian sheep

Methane Phenotype Approaches

1 Group 1

Whole-signal summary

Summarize all values together

Mean | Median | Area under curve (AUC)

Fast and reproducible, but averages breathing, eructation bursts, and background together.

2 Group 2

Peak detection + summary

Identify high concentration events only

Local peaks | Peak algorithms

AMPD | Local maxima | Alternative peak approach

Focuses on moments of elevated concentration using different peak-detection algorithms.

3 Group 3

Separate respiration and eructation

Classify peaks and report separate phenotypes

Eructation phenotype

Respiration phenotype

Threshold methods

Mean + SD rule | IQR (boxplot) rule | Statistical mixture model

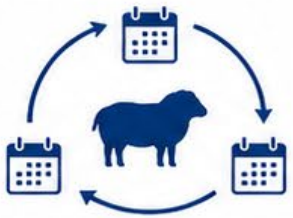
Most biologically explicit approach: derives distinct phenotypes for breathing and eructation components.

Different analytical methodologies applied to the same raw LMD data can produce very different methane phenotypes and completely different animal rankings.

The Evaluation Framework: Four Criteria

We evaluate methane phenotype approaches using four complementary criteria to determine which methods are **"better."**

1 Repeatability



If you measure the same animal three days in a row, does the method give consistent results?

How we measure it

Estimated as the proportion of total variance explained by the animal (r)

$$r = \frac{Var_{Animal}}{Var_{Total}}$$

High repeatability = consistent measurements for the same animal

2 Biological sensitivity

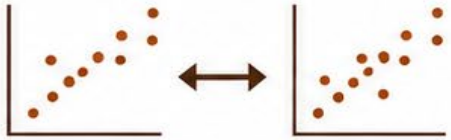
Does the method detect known drivers of methane variation — feeding time, body weight, diurnal period, breed?

How we measure it

Tested by likelihood ratio tests in a mixed model.

High biological sensitivity = captures real biological drivers of methane

3 Between-method agreement



Do different methods agree on the absolute value of the phenotype?

How we measure it

Measured by concordance correlation coefficient (CCC), which penalises both poor correlation and scale differences.

High agreement = methods produce similar absolute values

4 Animal ranking consistency

Do different methods agree on which animals are high vs low emitters?

How we measure it

Measured by Spearman rank correlation.

High consistency = consistent animal ranking; most breeding-relevant

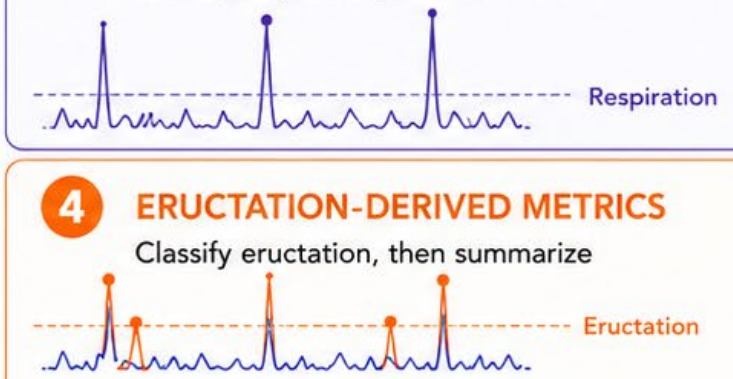
A better method scores high on all four criteria: consistent, biologically meaningful, agreeable with other methods, and reliable for ranking animals.

RESULTS

THE METHANE PHENOTYPES WE DERIVED

From raw LMD signals, we extracted 13 phenotypes using 7 analytical approaches.

APPROACH GROUP	PHENOTYPE	WHAT IT MEANS
1 WHOLE-SIGNAL METRICS Summarize all values together	CH_4/day	Estimated CH_4 production, g/day
	WS-AUC	Area under methane curve per minute
	WS-Mean	Mean methane concentration using all values
2 PEAK-BASED METRICS Identify peaks, then summarize	Peak-S	Mean of simple local peaks
	Peak-A	Mean of AMPD-detected peaks
3 RESPIRATION-DERIVED METRICS Classify respiration, then summarize	Resp-SD	Respiration mean using mean + 1 SD threshold
	Resp-IQR	Respiration mean using IQR threshold
	Resp-GMM	Respiration mean from Gaussian mixture model
	Resp-AG	Respiration mean from AMPD + mixture model
4 ERUCTATION-DERIVED METRICS Classify eructation, then summarize	Eruct-SD	Eructation mean using mean + 1 SD threshold
	Eruct-IQR	Eructation mean using IQR threshold
	Eruct-GMM	Eructation mean from Gaussian mixture model
	Eruct-AG	Eructation mean from AMPD + mixture model

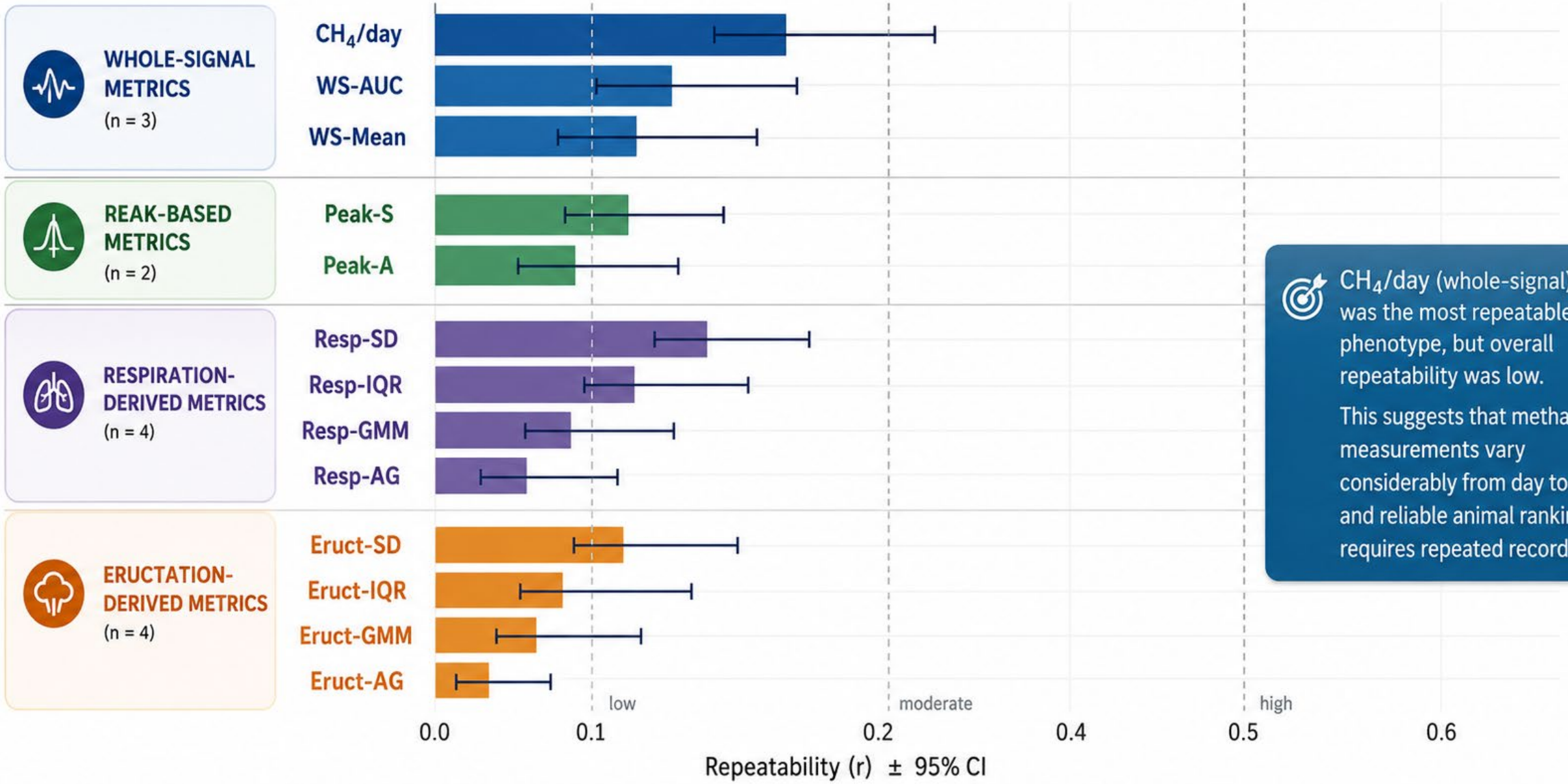


Respiration peaks Eructation peaks

WS = whole-signal; Peak-S = simple peaks; Peak-A = AMPD peaks; Resp = respiration; Eruct = eructation; SD = mean + 1 SD threshold; IQR = interquartile range threshold; GMM = Gaussian mixture model; AG = AMPD + Gaussian mixture model.

Day-to-day repeatability by phenotype

r = VarAnimal / VarTotal (LMM: Day + Site as fixed effects, Animal as random)



CH_4/day (whole-signal) was the most repeatable phenotype, but overall repeatability was low.

This suggests that methane measurements vary considerably from day to day, and reliable animal ranking requires repeated records.

Sensitivity of methane phenotypes to biological and management drivers

Likelihood-ratio tests from linear mixed models

	PHENOTYPE	DRIVERS TESTED: Sex	Weight	Feed time	Diurnal period	Breed	Day
WHOLE-SIGNAL METRICS (n = 3)	CH_4/day	*	ns	ns	***	***	*
	WS-AUC	*	ns	ns	***	***	*
	WS-Mean	*	ns	ns	***	***	*
PEAK-BASED METRICS (n = 2)	Peak-S	*	*	ns	*	*	ns
	Peak-A	ns	***	ns	ns	ns	*
RESPIRATION-DERIVED METRICS (n = 4)	Resp-SD	*	*	*	***	***	***
	Resp-IQR	ns	***	*	***	***	***
	Resp-GMM	ns	***	*	*	***	***
	Resp-AG	ns	ns	ns	*	*	ns
ERUCTATION-DERIVED METRICS (n = 3)	Eruct-SD	ns	ns	ns	ns	*	ns
	Eruct-IQR	ns	***	*	ns	ns	ns
	Eruct-GMM	ns	*	ns	ns	*	ns

***	$p < 0.001$
*	$p < 0.05$
ns	Not significant
NA	Not estimable

KEY MESSAGE

Respiration phenotypes carry biological signal; eructation phenotypes mostly carry noise.

Respiration-based phenotypes like Resp-SD, Resp-IQR, Resp-GMM, and Resp-AG detect all six known drivers: sex, weight, feeding time, diurnal period, breed, and recording day.

Animal rank stability across methane phenotyping methods

Spearman rank correlation of animal rankings (animal-level means)

		CH_4/day	WS-AUC	WS-Mean	Peak-S	Peak-A	Resp-SD	Resp-IQR	Resp-GMM	Resp-AG	Eruct-SD	Eruct-IQR	Eruct-GMM
Whole-signal (n = 3)	CH_4/day	1.00											
	WS-AUC	0.86	1.00										
	WS-Mean	0.86	0.61	1.00									
Peak-based (n = 2)	Peak-S	0.86	0.61	0.83	1.00								
	Peak-A	0.61	0.61	0.83	0.57	1.00							
Respiration-derived (n = 4)	Resp-SD	0.71	0.58	0.72	0.94	0.21	1.00						
	Resp-IQR	0.58	0.58	0.69	0.25	0.90	0.00	1.00					
	Resp-GMM	0.71	0.58	0.38	0.94	0.00	0.08	0.08	1.00				
	Resp-AG	0.39	0.39	0.63	0.10	0.86	−0.10	0.84	0.11	1.00			
Eructation-derived (n = 3)	Eruct-SD	0.82	0.47	0.67	0.84	0.12	0.83	0.11	0.03	0.30	1.00		
	Eruct-IQR	0.47	0.47	0.80	0.21	0.74	0.03	0.74	0.07	0.68	0.57	1.00	
	Eruct-GMM	0.37	0.37	0.61	0.10	0.71	−0.07	0.70	0.01	0.82	0.60	0.60	1.00

Spearman rank correlation: 1.00, 0.50, 0.00, −0.10

Whole-signal (n = 3) | Peak-based (n = 2) | Respiration-derived (n = 4) | Eructation-derived (n = 3)

Key messages

Which animal you select depends almost entirely on which method you use, not on the animal's biology. Two-thirds of all methodology pairs rank animals so differently.

within-family correlations are generally strong (high blues), while cross-family comparisons are weaker or near zero (green/yellow).

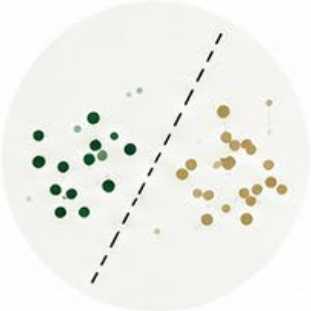
THE LMD WORKS — BUT ONLY IF WE AGREE ON HOW TO USE IT.

The LMD works.

Fifteen years of research have produced a device that reliably captures methane from sheep and goats in field conditions.

The signal is real.

Respiration-based classification methods detect sex, body weight, feeding time, diurnal period, and site effects consistently. The biology is there.



But results depend on the method.

Different analytical choices applied to the same recording produce phenotypes that rank animals differently two-thirds of the time.

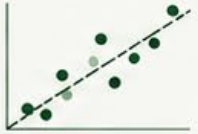
This study was unable to recommend one phenotype to be used for future analysis. However, with a larger dataset and more extensive analysis, we may come to a better conclusion.

THE PATH FORWARD IS CLEAR

1. **Standardize** around respiration-peak classification.

2. **Require** a minimum number of repetitions.

3. **Validate** against a reference method.



Once those three steps are taken, LMD becomes a **viable large-scale phenotyping tool** for methane breeding programs in small ruminants — including in resource-limited field settings where no other method is practical.

THE INSTRUMENT IS READY. THE PROTOCOL JUST NEEDS TO CATCH UP.