animals

MDPI

Article

Exploratory Evaluation of Agreement and Diagnostic Accuracy of a Laser Methane Detector Relative to GreenFeed Measurements in Spanish Boer Goats

Hossam Ismael, Mariline Hilaire, Santosh Chaudhary, Vivian Kanyi, Heba Abdo, Nar Gurung, Bipana Budha and Byeng R. Min *

Department of Agricultural and Environmental Sciences, Tuskegee University, Tuskegee, AL 36088, USA; hismael@tuskegee.edu (H.I.); mhilaire6334@tuskegee.edu (M.H.); schaudhary8992@tuskegee.edu (S.C.); vkanyi1503@tuskegee.edu (V.K.); habdo@tuskegee.edu (H.A.); bbudha6391@tuskegee.edu (B.B.)
* Correspondence: bmin1@tuskegee.edu

Simple Summary

Methane produced by goats and other ruminant animals contributes to greenhouse gas emissions, so reliable and practical methods are needed to measure these emissions. This study compared two methods for measuring methane from Spanish Boer goats: a portable laser-based detector and an automated feeding-based measurement system. Twenty-six goats with matched measurements were included in this exploratory, retrospective comparison. The results showed that measurements from the two methods did not closely agree, and the portable laser method could not reliably identify goats with higher or lower methane emissions compared with the automated system under the conditions of this study. However, the relatively small number of animals limited the ability of the study to detect weaker relationships between the two methods. In addition, the two technologies measure methane differently, and the automated system itself has not been fully established as a reference method for goats. Overall, the findings indicate that the portable laser detector cannot yet be recommended as a reliable replacement or screening tool for measuring methane emissions in goats. Larger studies using more animals, different locations and conditions, standardized measurement procedures, and comparison with established reference methods are needed before its usefulness for goats can be determined with confidence.

Abstract

Accurate quantification of enteric methane emissions from ruminants is needed for greenhouse gas inventories, mitigation research, and genetic selection. The laser methane detector (LMD) is a portable, low-cost tool that measures methane concentration along a laser path, whereas the GreenFeed (GF) system estimates methane flux; validation data for small ruminants are scarce. This exploratory, retrospective analysis evaluated previously collected paired LMD and GF measurements from 26 Spanish Boer goats (137 observations; mean 5.3 per animal) at Tuskegee University. GF was treated as a comparator rather than a validated reference. Repeated-measures correlation, cluster-robust regression, repeated-measures Bland–Altman analysis, and observation-level diagnostic accuracy at two thresholds were used. No detectable relationship was observed between LMD and GF: the repeated-measures correlation was r = 0.012 (p = 0.90), and the regression slope was 0.018 ± 0.065 (p = 0.78; cluster-bootstrap 95% CI −0.115 to 0.138). The Bland–Altman analysis showed systematic numerical differences between the two instruments (mean difference −5.54 g CH_4/day; limits of agreement −32.93 to 21.86), but because LMD (ppm·m) and GF (g/day) are expressed in different physical units, this was not interpreted as underestimation by the LMD. Observation-level discrimination was close to chance (AUC = 0.544,

Check for updates

Academic Editor: Guoming Li

Received: 9 November 2026
Revised: 20 November 2026
Accepted: 23 November 2026
Published: 29 November 2026

Copyright: © 2026 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license.

95% CI 0.459–0.636; $\kappa = -0.022$ at 13 g CH_4/day). With 26 animals, only a large association ($|r| \geq 0.53$) could have been reliably detected, and GF visit-level quality information was unavailable. The findings are therefore exploratory, with limited precision, and additional validation, ideally against respiration chambers, is required before the LMD can be recommended for methane phenotyping in goats.

Keywords: diagnostic accuracy; enteric fermentation; goats; GreenFeed system; greenhouse gases; laser methane detector; methane emissions; method comparison; repeated measures

1. Introduction

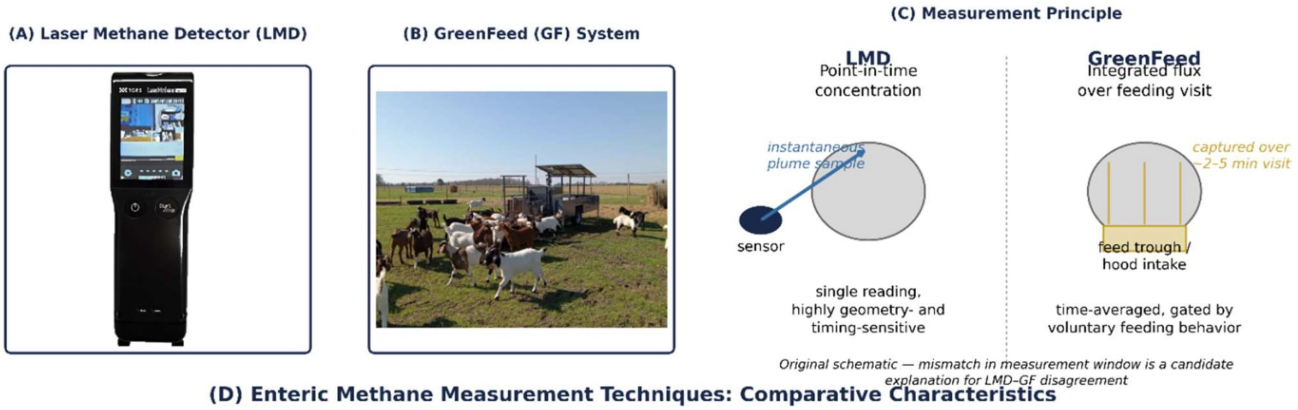
Methane (CH_4) has a global warming potential substantially greater than that of carbon dioxide (CO_2) over a 100-year horizon, and enteric CH_4 from ruminant livestock is among the largest contributors to agricultural greenhouse gas (GHG) emissions worldwide [1,2]. Accurate measurement of these emissions is essential for national GHG inventories, evaluating mitigation strategies, and genetic selection programs to reduce emission intensity [3]. CH_4 output is also partly under host genetic control, making individual-animal phenotyping at scale a prerequisite for selection [4,5]. The scale of livestock CH_4 reductions needed to support stringent temperature goals has been critically assessed [6].

Respiration chambers and the sulfur hexafluoride (SF_6) tracer technique are generally regarded as reference, or "gold-standard," methods because they can directly quantify gas exchange; however, their considerable cost, labor demands, and inherently low throughput limit their use to small numbers of animals [7,8]. Moreover, direct SF_6–chamber comparisons in dairy cows show that estimates depend on the technique used [9]. This limitation has, in turn, driven considerable interest in field-deployable alternatives, notably the GF automated head-chamber system, breath "sniffers," portable accumulation chambers, and laser-based CH_4 detectors [10–13]; non-invasive individual-animal measurement in dairy cows has been reviewed [14].

The GF system (C-Lock Inc., Rapid City, SD, USA) estimates CH_4 production (g CH_4/day) by combining continuously monitored gas concentrations with airflow measurements, as animals voluntarily visit a baited feeding station [15]. A range of validation studies report reasonably satisfactory agreement between GF and either respiration-chamber or SF_6-based estimates; notably, GF repeatability improves substantially as measurement periods lengthen, supporting its adoption in genetic-evaluation programs [15–18]. In cattle, GF measurements also agree closely with predictions from empirical equations derived largely from respiration-chamber data (83 treatment means; R^2 mostly > 0.90), and in direct comparisons with chambers (n = 20 treatment means) the two methods were closely related ($R^2 = 0.92$), with GF values numerically about 13 g/day higher than chamber values [19]. Because GF relies on short spot measurements, longer measurement periods and larger numbers of animals are needed than with chambers to offset the greater variability of individual spot samples [19]. Concordance between GF and other methods nevertheless varies among field comparisons in dairy cows and beef heifers, partly because feed intake differs between chambers and field conditions [19–21]. Its principal practical drawbacks, nonetheless, remain equipment cost and dependence on voluntary animal visits, which can considerably bias sampling when visit frequency varies across the day [15,18]. Evidence on GF in goats is considerably thinner. In confinement, a GF unit used with goats returned CH_4 production of 25.6–27.3 g/day and yields of about 42–44 g/kg dry matter (DM) intake, roughly 2.3- to 2.5-fold and 2.2- to 2.3-fold the 11.1 g/day and 19.3 g/kg DM intake obtained with head-box calorimetry, although GF and calorimetry values were moderately to strongly correlated across animals (r = 0.62–0.70 for g/day) [22]. In grazing

goats, daily CH_4 production by GF (about 26 g/day) did not differ with the number of goats per unit, whereas visit frequency and CH_4 yield (g/kg DM intake) did, and estimates were more variable with 12 goats than with 18–24 goats per unit [23].

The Laser Methane Detector (LMD) employs tunable diode laser absorption spectroscopy, a technique originally developed for detecting industrial gas leaks [24] and first adapted for livestock CH_4 monitoring by Chagunda et al. [25], which was subsequently validated against respiration chambers in dairy cattle [26,27]. Portable, low-cost, and non-invasive, the LMD has since been used in cattle under both grazing and barn conditions [28–31], in sheep [32–34], and, much less commonly, in goats [35]. However, it is critical to note that the LMD measures CH_4 concentration along the laser path (ppm·m) rather than CH_4 flux (g/day); converting concentration to a flux-comparable estimate thus requires a calibration or prediction model, absent which the two instruments are, quite simply, measuring physically different quantities (Figure 1; [27,36]). LMD readings are also well documented as sensitive to wind speed, humidity, animal movement, measurement distance, and operator technique [25,37].



(D) Enteric Methane Measurement Techniques: Comparative Characteristics

Method	Setting	Gas Measured	Digestive Origin	Measurement Mode	Output Metric
Respiration chambers	Indoor	CH_4, multi-gas	Rumen, hindgut	Continuous	Flux
Sulphur hexafluoride tracer	Indoor, grazing	CH_4	Rumen	Continuous	Flux
GreenFeed	Indoor, grazing	CH_4, multi-gas	Rumen	Short-term	Flux
Sniffer method	Indoor	CH_4, multi-gas	Rumen	Short-term	Concentration
Ventilated hood	Indoor	CH_4, multi-gas	Rumen	Continuous	Flux
Facemask	Indoor	CH_4, multi-gas	Rumen	Short-term	Flux
Laser methane detector	Indoor, grazing	CH_4	Rumen	Short-term	Concentration
Portable accumulation chamber	Indoor	CH_4, multi-gas	Rumen, hindgut	Short-term	Concentration

Figure 1. Overview of the enteric methane (CH_4) measurement approach and the two technologies (Laser methane detector (LMD) vs. GreenFeed (GF) system compared in this study.

Among comparisons of LMD with GF, the study by Sorg et al. [36] in dairy cows using two breath analyzers is arguably the most influential. It reported a concordance correlation coefficient (CCC) of 0.02 and a coefficient of individual agreement (CIA) of 0.06 between raw LMD and GF values, whereas the repeated-measures correlation was considerably higher at 0.66. After applying a calibration/prediction equation to convert LMD readings into an estimated flux, the CCC rose to 0.22 and the CIA to 0.99. This pattern of poor raw agreement, yet markedly improved agreement following calibration, together with a reasonable capacity to rank animals, recurs consistently across the LMD validation

literature in both cattle and sheep [26,28,32,38,39], and motivates two design principles that we consider essential for any new LMD-versus-GF comparison: (i) LMD values ought to be calibrated to a flux-comparable unit prior to any numerical comparison against GF output, and (ii) inasmuch as animals typically contribute multiple paired observations, the statistical analysis must account for within-animal clustering rather than treating all observations as though independent [40–44].

Comparative work conducted specifically in goats, however, remains scarce: the majority of LMD validation studies to date have been conducted in dairy cattle [25–29,36] or sheep [32,33,45], with only a small number applying the LMD to goats under farm conditions [35]. Because small ruminants differ appreciably from cattle in body size, breath-plume characteristics, and typical measurement geometry, agreement statistics established for cattle cannot be assumed to apply to goats. Likewise, GF performance in goats rests on only two recent studies from a single research group [22,23], so any LMD–GF comparison in this species is made against a comparator that is still being characterized.

The present study has two objectives. It aims to (1) quantify agreement between LMD and GF measurements in Spanish Boer goats under practical farm conditions, using continuous-agreement and diagnostic-accuracy frameworks appropriate for clustered, repeated-measures data, and (2) evaluate whether LMD, with or without calibration, can serve as a practical alternative or complementary screening tool to GF for CH_4 phenotyping in this species. To our knowledge, this is among the first studies to compare these two technologies in a small-ruminant species other than sheep, and it aims to provide a foundation for future calibration-model development in goats. Given this, GF is treated throughout as the comparator method rather than a validated absolute reference.

We hypothesized that LMD measurements would demonstrate sufficient association and classification performance relative to GreenFeed measurements to support their use as a practical methane phenotyping tool in goats. Statistical evaluation of this hypothesis was conducted using correlation, regression, agreement, and diagnostic accuracy analyses. This study is a retrospective analysis of previously collected paired measurements rather than a purpose-designed validation trial. Given the uncertainty surrounding GreenFeed validation in goats, the present study should be regarded as an exploratory comparator analysis rather than a definitive validation study. The potential significance of this work is fourfold. First, direct comparisons between LMD and GF systems in goats are extremely limited. Second, the existing LMD validation literature has focused primarily on cattle and sheep, so agreement statistics from those species cannot be assumed to apply to goats. Third, this study provides one of the first exploratory evaluations of agreement and diagnostic performance between these two technologies in Spanish Boer goats. Finally, the findings establish a foundation for future calibration-model development and for more rigorous validation studies in goats that incorporate independent reference methods.

2. Materials and Methods

The Tuskegee University Animal Care and Use Committee approved the procedures for animal care, handling, and sampling for this experiment (#R08-2024-06).

2.1. Animals and Housing

This study was a retrospective analysis of previously collected paired LMD and GF measurements rather than a purpose-designed, prospective validation experiment. The paired LMD–GF dataset analyzed here has not been previously published. This study used Spanish Boer goats aged 4–6 months, housed at the Tuskegee University Research and Education Unit. Forty-five animals were originally enrolled and monitored using LMD (identified in the raw data by a "Normal Tag" number). Body weight, Body Condition Score

(BCS), and FAMACHA anemia score were available for 25 of the 26 animals ultimately retained in the paired analysis, courtesy of routine herd health monitoring (Table 1); notably, sex was not recorded in any of the source data reviewed for this study. GF continuous-monitoring data (identified by a considerably longer RFID "Magnetic Tag" number) were recorded for 33 of the 45 animals over a 19-day window (10–28 March 2025), with LMD spot measurements undertaken on 9 of those days. A tag crosswalk table linking Normal Tags to their corresponding Magnetic Tags was used to match LMD and GF records by animal and date; this crosswalk covered 40 of the 45 enrolled animals, and, of the 33 GF-monitored animals, complete matched LMD–GF records across at least one common date were ultimately available for 26. The analyzed sample for the present study is therefore 26 goats, contributing 137 paired LMD–GF observations (mean 5.3 observations per animal, range 5–6); the remaining 19 enrolled animals lacked either GF coverage, a resolvable tag linkage, or a temporally matching LMD reading, and were accordingly excluded from the agreement analysis. This attrition, by no means uncommon in observational field studies that depend upon voluntary GF visits and multi-system tag linkage, is treated herein as a limitation of the achieved sample size (Section 4) rather than as a gap to be retrospectively filled; all statistics reported in Section 3, together with this manuscript's abstract, tables, and conclusions, refer consistently to this 26-animal, 137-observation sample. No missing values were imputed, estimated, or reconstructed; analyses were restricted to animals with complete paired LMD–GF records.

Table 1. Descriptive summary of the analyzed sample and raw measurements.

Characteristic	Value
Animals enrolled (LMD-monitored)	45
Animals with GF coverage	33
Animals with complete matched LMD–GF records (analyzed sample)	26
Total paired observations	137
Observations per animal, mean (range)	5.3 (5–6)
Body weight, mean ± SD (range) [n = 25/26; units as recorded, not specified in source file]	32.94 ± 2.85 (28.0–40.5)
Body Condition Score (BCS, 1–5 scale), mean ± SD (range) [n = 25/26]	2.56 ± 0.33 (2.0–3.5)
FAMACHA anemia score (1 = best–5 = worst), mean ± SD (range) [n = 25/26]	2.12 ± 0.93 (1–4)
Sex ratio	Not recorded in available source files
LMD reading (g-equivalent units), mean ± SD (range)	12.29 ± 9.89 (2.00–51.00)
LMD reading, median	9.00
GF CH_4, g/day, mean ± SD (range)	17.83 ± 10.03 (2.21–50.91)
GF CH_4, g/day, median	15.27
Measurement window	10–28 March 2025 (9 LMD measurement days)
Ambient temperature during measurement days, °F (n = 112 obs. with weather match)	73.5 ± 6.0 (61.7–81.3)
Wind speed during measurement days, mph (n = 112)	30.8 ± 23.4 (4.4–69.8)

Two auxiliary datasets from the same herd were reviewed but are not part of the paired LMD–GF comparison reported herein: a separate 3-day GF-only repeatability trial and a

distinct biochar additive trial employing LMD alone. These datasets constitute separate studies and were not incorporated into the current device-comparison analysis.

Animals were not randomly allocated to treatment groups, as the study was observational (both instruments were applied to the same animals); blinding was not applicable to instrument operators, a limitation noted in Section 4. We did not perform an a priori sample-size or power calculation; instead, the analyzed sample of 26 animals reflects those with complete, matched LMD–GF records, and Sections 2.5 and 4 report the corresponding achieved statistical precision.

All procedures were conducted in accordance with national and institutional regulations governing the care and use of agricultural animals in research, and followed the ARRIVE 2.0 guidelines for reporting animal research [46,47], to the extent applicable to an observational instrument-comparison study of this nature.

2.2. Green Feed (GF) System Measurements

CH_4 emissions were measured using a GF Pasture System (C-Lock Inc., Rapid City, SD, USA), which served as the comparator method rather than a validated gold-standard reference for goats. Goats were attracted to the unit by offering small quantities of feed during each visit (40 g). All 33 GF-monitored goats shared this single unit, a group size that satisfies the minimum of 18 goats per unit recommended for grazing goats (18–24 goats) and lies within the approximately 20–50 animals per unit suggested by the manufacturer [23]. The system continuously monitored CH_4, carbon dioxide, oxygen, and hydrogen concentrations, along with airflow rate, and used the manufacturer's proprietary algorithm to estimate CH_4 production (g CH_4/day) for each visit. The GF values used in the present analysis were the system's 1-day rolling-average CH_4 production (g/day) per animal, computed by the GF software and exported directly from the unit; the minimum visit duration and number of qualifying visits underlying each daily rolling average are, it should be noted, internal to the GF algorithm and were not separately exported alongside the averaged values (visit counts per animal-day, where recorded, ranged from 1 to 10 in the exported event-count log). Because published guidance for GF calls for at least two valid visits per animal per day and 30–50 visits per animal over the measurement period [22,23], animal-days with a single recorded visit fall below the recommended minimum, and the exported averages did not allow screening animal-days against this criterion. For GF unit operation, we imposed no fixed visitation schedule. However, we restricted access to a maximum of 10 visits per animal per day, with a feed allocation of 40 g of pellets per visit. We conducted daily gas calibration and CO_2 recovery assessments to ensure measurement accuracy and system performance. We collected data from all goats enrolled in the study that voluntarily entered the GF unit in response to the system-generated auditory cue or the sound of pellets being dispensed into the feed trough [14–16]. Because visit-level quality information (such as visit duration and the number of valid visits per animal-day) was unavailable, and some animal-days had only a single recorded visit, GF output is treated throughout as a comparator method of uncertain precision rather than as a validated gold-standard reference.

2.3. Laser Methane (CH_4) Detector (LMD) Measurements

We measured CH_4 concentration using a handheld LMD employing infrared tunable diode laser absorption spectroscopy, directed at the goat's muzzle/exhaled breath at a fixed distance. We measured CH4 emissions using a Laser CH_4 SMART LMD (Gastar Co., Ltd., Tokyo, Japan), a portable infrared laser-based device designed for remote CH_4 detection. We positioned the detector approximately 50 cm from the animal's nostrils during measurement. We monitored each animal for 60 s per sampling event, with measurements conducted four times daily at predetermined intervals. Two trained operators performed the

measurements. The Laser CH_4 SMART detects CH_4 remotely by utilizing an infrared laser and reports CH_4 concentration as ppm·m (gas concentration × optical path length). Daily weather conditions (temperature, rainfall, and wind speed) at the study site were recorded independently by means of the on-site climatological record, and were merged with the paired dataset by measurement date for the exploratory analysis presented in Section 3.5; wind speed on several measurement dates was, notably, unusually high (reaching up to 69.8 mph), a consideration of some relevance given the LMD's well-documented sensitivity to wind [25,37].

2.4. Unit Conversion and Calibration of Laser Methane Detector (LMD) Readings

The GF reports CH_4 production as mass flux (g CH_4/day), whereas the LMD records CH_4 concentration integrated over the laser path (ppm·m). Because these are physically different quantities, prior work indicates that raw, unconverted LMD values show poor numerical agreement with GF or respiration-chamber flux estimates, even when the two instruments rank animals comparably [36,38]. A mass-conversion relationship (Mass (g) = PPM × Volume (m^3)/1000) has been employed elsewhere as the basis for a prediction equation relating LMD output to GF-comparable CH4 mass, following the approach adopted by Sorg et al. [36].

No path-length/volume field or separate conversion factor accompanied the raw LMD readings analyzed herein (integer range 1–51 across the dataset); therefore, we compared raw LMD readings directly against GF g/day values in Section 3. In place of a physical unit conversion, we fitted an empirical calibration/prediction equation by regressing GF on LMD (following Sorg et al. [36]) using the cluster-robust regression described in Section 2.5. The resulting relationship (GF = 17.6 + 0.018 × LMD, cluster-robust SE = 0.065) has a slope statistically indistinguishable from zero; as such—in contrast to Sorg et al. [36], where calibration substantially improved agreement—no useful GF-predictive information could be extracted from LMD in the present dataset. This finding ought to be interpreted alongside the achieved statistical precision reported in Sections 2.5 and 3.2: with only 26 animals, the study could reliably detect only a fairly large true LMD–GF association, and as such this result does not rule out a smaller true relationship that a larger sample might yet detect (Section 4).

2.5. Statistical Analysis

Continuous agreement between LMD and GF measurements was assessed using pooled Pearson and Spearman correlation coefficients, linear regression, and Bland–Altman analysis [40], while diagnostic accuracy was assessed by dichotomizing GF values and computing sensitivity, specificity, positive and negative predictive values, overall and balanced accuracy, Cohen's κ [48,49], McNemar's test [50], and the area under the receiver operating characteristic curve (AUC) [51]. Because the 137 observations are clustered within 26 animals (Section 2.1), all of the above were implemented using methods appropriate for repeated measures and a modest number of clusters, namely:

Repeated-measures correlation. In addition to the naive pooled Pearson/Spearman coefficients, we report a within-animal (repeated-measures) correlation, computed by mean-centering both LMD and GF values within each animal and correlating the centered values—a procedure equivalent to partialing out animal identity [41,44].

Cluster-robust regression with a small-sample correction. GF was regressed on LMD using ordinary least squares with a cluster-robust ("sandwich") variance estimator clustered by animal, including the standard finite-sample correction factor $G/(G-1) \times (n-1)/(n-k)$. Inference was based on the t-distribution with $G-1$ degrees of freedom ($G = 26$ clusters), rather than the normal approximation—the conventional adjustment

when the number of clusters is modest [52]. Given that cluster-robust sandwich estimators can remain unstable with fewer than approximately 30–50 clusters, every cluster-robust point estimate was cross-validated against a nonparametric cluster (pairs) bootstrap (item 5, below); the two approaches agreed closely for the key regression slope (analytic 95% CI −0.110 to 0.146 vs. bootstrap 95% CI −0.115 to 0.138), a finding we take as evidence that the small number of clusters is not materially distorting inference in the present case.

Repeated-measures Bland–Altman. We computed limits of agreement using the repeated-measures extension of Bland and Altman [42,43], which partitions the variance of the LMD–GF differences into between-animal and within-animal components before computing the 95% limits, rather than treating the 137 differences as independent.

Proportional bias test. We regressed the difference (LMD − GF) on the mean of the two measurements, again with cluster-robust standard errors, to formally test for any magnitude-dependent bias pattern.

Cluster bootstrap confidence intervals. The 95% confidence intervals reported in Section 3 (mean bias, Pearson r, regression slope, AUC) were computed by resampling animals—rather than individual observations with replacement across 2000–3000 replicates and taking the 2.5th and 97.5th percentiles; this is the standard, generally more robust, approach to propagating clustering uncertainty into interval estimates when the number of clusters is small [52].

Diagnostic threshold. Diagnostic accuracy is reported at two thresholds: 13 g CH_4/day, a round-number cutoff with some precedent in prior LMD field-screening discussions, albeit without farm-specific clinical justification for the present dataset; and a Youden-optimal threshold (the point on the ROC curve maximizing sensitivity + specificity − 1) [53], derived from an above/below-median GF split. We emphasize that this median-split outcome definition constitutes a data-driven convenience for threshold selection rather than a biologically or operationally validated definition of a "high-emitting" goat, and that both thresholds ought to be read as a sensitivity analysis on threshold placement rather than as a validated diagnostic cutoff; a farm- or program-specific definition of a CH4-positive animal (e.g., a selection-index percentile) would be required before either threshold could be applied operationally. The ROC and diagnostic-accuracy analyses were performed at the observation level (individual paired observations), not at the animal level, and are therefore exploratory; they do not constitute animal-level methane phenotyping. Confusion matrices for both thresholds are provided in Table 2.

Table 2. Observation-level confusion matrices for LMD classification relative to GF comparator classification (137 paired observations; GF-positive and GF-negative as defined in Section 2.5).

Threshold: 13 g CH_4/Day	GF Positive	GF Negative
LMDpositive	TP = 24	FP = 19
LMD negative	FN = 55	TN = 39
Youden-optimal threshold (LMD ≥ 8; above/below-median GF split):		
Youden-optimal threshold	**GF positive**	**GF negative**
LMD positive	TP = [to be completed from raw data]	FP = [to be completed from raw data]
LMD negative	FN = [to be completed from raw data]	TN = [to be completed from raw data]

Statistical power. Because the analyzed sample comprises only 26 independent clusters, we report the minimum true repeated-measures correlation that the study had 80% power to detect at $\alpha = 0.05$ (Section 3.2) alongside the observed effect, so that a non-significant result might be correctly interpreted as "not detected, with limited precision" rather than as "shown to be absent."

Sensitivity and specificity are defined herein as the probability that the LMD correctly classifies a GF-positive observation as positive (sensitivity, true positive rate) and the probability that it correctly classifies a GF-negative observation as negative (specificity, true negative rate), each performed in Python 3 (pandas, numpy, scipy, scikit-learn). We used a two-sided $\alpha = 0.05$ throughout.relative to the GF comparator. Analyses were performed in Python 3 (pandas, numpy, scipy, scikit-learn). We used a two-sided $\alpha = 0.05$ throughout.

3. Results

Results are presented sequentially by objective: sample description (Section 3.1), agreement and calibration (Section 3.2), numerical differences between instruments (Section 3.3), and diagnostic performance (Section 3.4). All analyses are exploratory.

3.1. Sample Description

Body weight, BCS, and FAMACHA© scores were recorded during routine herd health monitoring for 25 of the 26 analyzed animals (one analyzed animal, Normal Tag 769, was absent from the health-monitoring roster, and its body-condition data are accordingly unavailable); mean FAMACHA© scores falling within the 1–3 range across the sample are consistent with a generally low burden of anemia/parasitism in this herd at the time of monitoring. Sex, it bears repeating, was not recorded in any of the source files reviewed, and is reported here as unavailable rather than assumed. All results that follow are computed using the repeated-measures-appropriate methods described in Section 2.5 on this 26-animal, 137-observation sample.

3.2. Agreement and Calibration (Correlation and Regression)

The pooled Pearson correlation was r = 0.018 ($p = 0.84$), and Spearman ρ was 0.104 ($p = 0.23$)—both negligible in magnitude. After accounting for repeated measures, the within-animal (repeated-measures) correlation was r = 0.012 (df = 110, $p = 0.90$; cluster-bootstrap 95% CI for pooled Pearson r: −0.108 to 0.143). Clustering did not mask a stronger underlying relationship, as the near-zero association holds at both the pooled and within-animal levels. With 26 clusters, the study had 80% power ($\alpha = 0.05$) to detect a repeated-measures correlation of $|r| \geq 0.53$. The observed r = 0.012, together with its confidence interval, falls comfortably within the range the study could resolve, and this result is therefore appropriately interpreted as no detectable association given the achieved sample—rather than as ruling out any true relationship weaker than 0.53 in magnitude.

The cluster-robust regression of GF on LMD yielded GF = 17.61 + 0.018 × LMD (slope SE = 0.065, 95% CI −0.110 to 0.146, $p = 0.78$; cluster-bootstrap 95% CI −0.115 to 0.138, closely mirroring the analytic interval), indicating no detectable linear relationship between LMD and GF in this dataset, at least at the achieved sample size.

3.3. Exploratory Repeated-Measures Bland–Altman Analysis

The mean bias (LMD—GF) was −5.54 g CH_4/day (cluster-bootstrap 95% CI: −8.14 to −2.94), indicating systematic numerical differences between the two instruments. Because LMD (ppm·m) and GF (g CH_4/day) are expressed in different physical units, this bias cannot be interpreted as evidence that the LMD underestimates methane production, and the Bland–Altman analysis is presented as exploratory. Partitioning the variance of the differences into within-animal (SD = 13.24) and between-animal (SD = 4.49) components yielded a total SD of 13.98 and repeated-measures 95% limits of agreement of −32.93 to 21.86 g CH_4/day (Figure 2)—limits that, notably, indicate poor interchangeability at the individual-observation level.

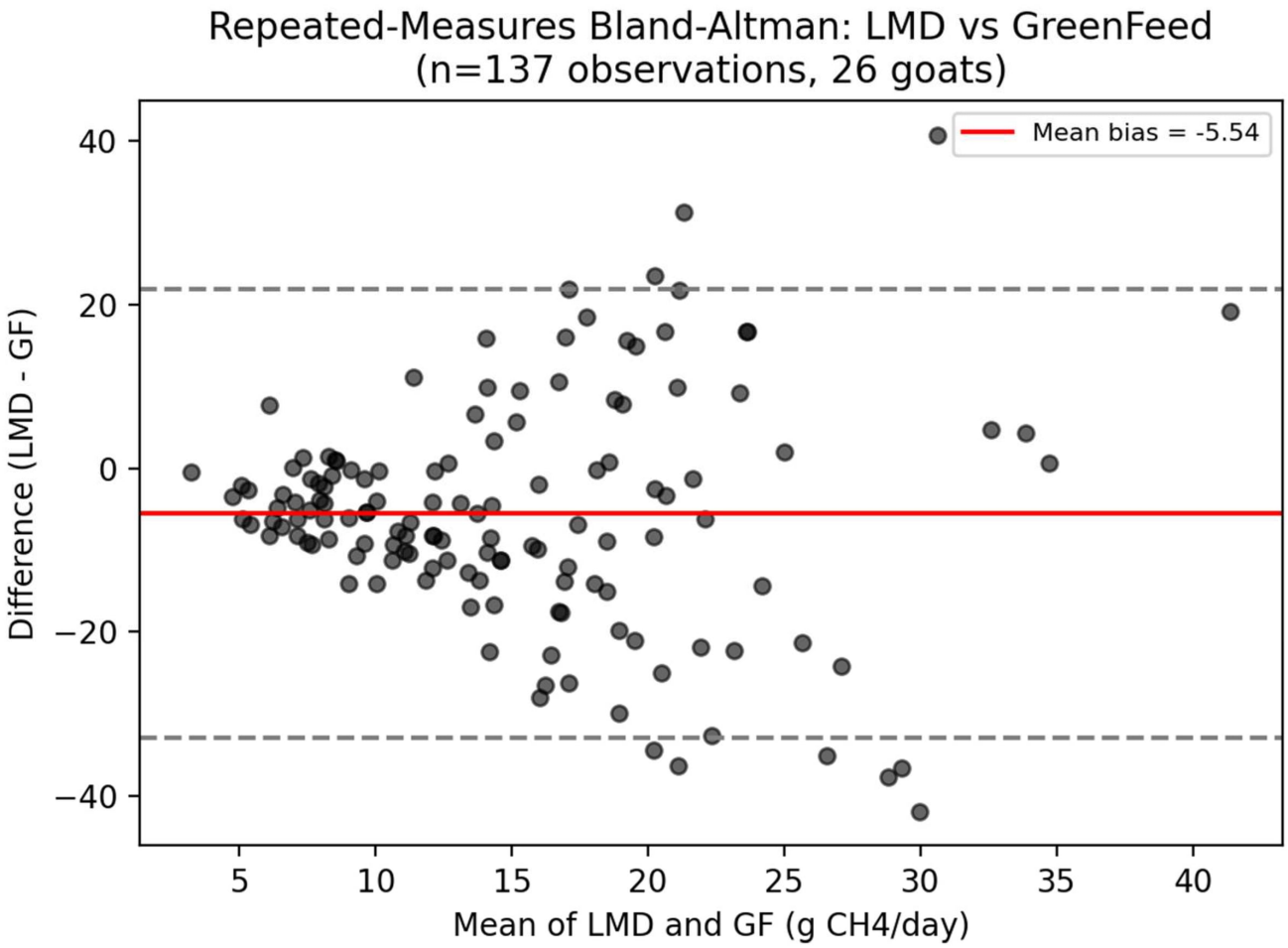


Figure 2. Repeated-measures Bland–Altman plot, Laser methane detector (LMD) vs. GreenFeed (GF) system (n = 137 observations, 26 goats). Solid red line = mean bias; dashed lines = 95% repeated-measures limits of agreement.

The proportional-bias test (difference regressed on mean, cluster-robust) was not statistically significant (slope = −0.027, cluster-robust SE = 0.236, $p = 0.91$; Figure 2): as such, the magnitude of LMD–GF disagreement does not appear to depend, in any detectable way, on the overall CH_4 level in this dataset.

3.4. Diagnostic Accuracy (Observation-Level, Exploratory)

Diagnostic accuracy is reported at two thresholds, treated here as a sensitivity analysis of threshold placement (Section 2.5, item 6), because neither has an independently validated clinical or operational basis in this dataset. At a threshold of 13 g CH_4/day: sensitivity = 30.4%, specificity = 67.2%, PPV = 55.8%, NPV = 41.5%, overall accuracy = 46.0%, balanced accuracy = 48.8%, Cohen's κ = −0.022, and McNemar's test $p < 0.0001$ (thereby confirming a significant asymmetry in the direction of misclassification). AUC for the underlying continuous classifier was 0.544 (cluster-bootstrap 95% CI: 0.459–0.636)—an interval that, tellingly, spans the chance value of 0.5. The ROC and diagnostic-accuracy analyses were performed at the observation level (individual paired observations), not at the animal level, and are therefore exploratory; they do not constitute animal-level methane phenotyping. Confusion matrices for both thresholds are given in Table 2.

At a Youden-optimal threshold (LMD ≥ 8, selected against an above/below-median GF split): sensitivity = 60.8%, specificity = 41.7%, balanced accuracy = 51.2%, κ = 0.010. Similarly, performance was close to chance at both thresholds, indicating that threshold

placement is not the limiting factor; rather, the underlying constraint is the weak, statistically undetectable LMD–GF association reported in Section 3.2, within the precision permitted by this sample size. Figure 3 illustrates the empirical ROC curve (AUC = 0.544).

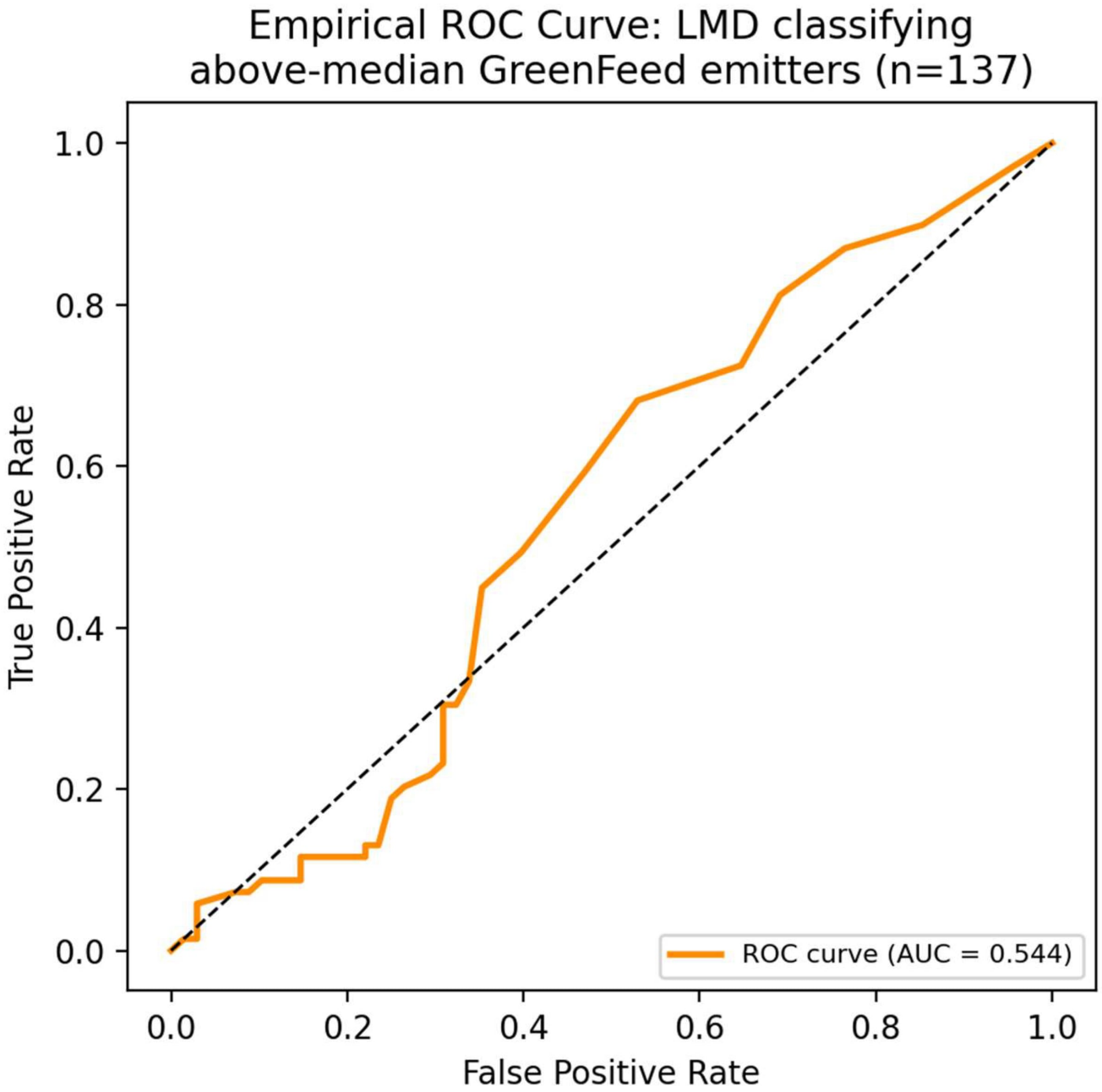


Figure 3. Empirical Receiver Operating Characteristic (ROC) curve for the Laser methane detector (LMD) classifying above- vs. below-median GreenFeed (GF) emitters (n = 137 observations; AUC = 0.544).

3.5. Exploratory: Environmental Covariates

Daily weather data (temperature, rainfall, wind speed) were available for most measurement dates in the on-site climatological record and were merged with the paired dataset accordingly. Wind speed on the measurement date was not significantly associated with the magnitude of the LMD–GF discrepancy (Pearson r = 0.102, n = 112, $p = 0.28$). However, wind speed was notably high on several dates (up to 69.8 mph on 15 March). Given the LMD's well-documented sensitivity to wind [25,37].

4. Discussion

In this exploratory, retrospective analysis of 26 Spanish Boer goats, no statistically detectable relationship was observed between LMD and GF measurements, whether assessed on a pooled or within-animal basis, and observation-level discrimination was close to chance at both thresholds. This is consistent, in direction, with dairy cattle studies reporting low concordance between raw, uncalibrated LMD and GF values [36,38], and agreement between field techniques and chambers varies with technique and protocol even for established methods [9,20,21]. Unlike Sorg et al. [36], who found that calibration substantially improved agreement, an empirical calibration here produced no improvement because the LMD–GF slope did not differ from zero. The Bland–Altman analysis showed systematic numerical differences between the instruments; however, because LMD (ppm·m) and GF (g/day) are different physical quantities, these differences should not be interpreted as underestimation of methane production by the LMD. No physical path-length/volume conversion was available, so a contribution of unit artifacts to the disagreement cannot be excluded and should be addressed in future work. The two instruments also sample fundamentally different quantities: the LMD gives an instantaneous, wind- and geometry-sensitive concentration reading [25,37], whereas GF integrates flux over voluntary visits [15,18], so the present results compare two field instruments rather than establishing the accuracy of either [14,31].

The reliability of GF as a comparator deserves explicit consideration, because the observed disagreement can arise from error in either instrument. Like the LMD, GF infers daily output from short spot measurements, an approach with precedent in sheep using portable chambers [54]. In cattle, GF concordance with chamber-based predictions is high for treatment means, but errors are expected to be larger for individual animal-days [19], which was our unit of analysis. In goats, GF estimates exceeded head-box calorimetry by more than two-fold (about 26 g/day versus 8.8–11.1 g/day) [22,23]; proposed explanations include a predominance of daytime visits (79%) and limited sensitivity of non-dispersive infrared sensors to the low breath CH_4 concentrations of small ruminants [22]. The mean GF value here (17.83 g CH_4/day) lies between these values, although differences in animals, diets, and protocols make this comparison only indicative. GF ranking of individual goats is also only moderately repeatable ($r = 0.20$–0.58 across group sizes) [23], which would attenuate any correlation attainable with another instrument, and estimates depend on visit frequency and group size [22,23]. All 33 goats here shared one unit, meeting the recommended minimum of 18 goats [23], but no fixed visit schedule was imposed, daily values were drawn from 1–10 visits per animal-day, visit-level quality information was unavailable, and we could not confirm that all animal-days met the recommended minimum of two valid visits [22,23]. GF was therefore treated as a comparator rather than a validated reference, and part of the LMD–GF disagreement may originate on the GF side; a three-way comparison including chamber or head-box calorimetry is needed to apportion the error [55,56].

Despite these results, the LMD's portability, low capital and operating cost, and non-invasive, real-time measurement continue to make it attractive for extensive and resource-limited production systems in which GF is impractical [10,13,36,39]. Under the conditions evaluated here, however, no statistically detectable relationship was observed, so we cannot support even a screening or ranking role for the LMD in goats without further validation; this contrasts with cattle studies in which LMD retains a reasonable capacity to rank animals despite poor absolute agreement [36,38]. Given the limitations of GF as a comparator, this finding should not be taken as evidence that the LMD is invalid against true emissions.

Several limitations warrant consideration. First, this was a retrospective analysis of previously collected data rather than a purpose-designed validation trial; animals were not randomly allocated and operators were not blinded. Second, the analyzed sample (26 of 45

enrolled animals, 137 observations) reflects attrition from incomplete GF coverage and tag-linkage gaps, and the study had 80% power to detect only a repeated-measures correlation of $|r| \geq 0.53$; a weaker true association could not have been reliably detected. Attrition and incomplete source records carry further consequences: the 19 excluded animals may have differed from the 26 retained animals, which limits generalizability and could affect the estimates; body weight, BCS, and FAMACHA records were missing for one analyzed animal and sex was not recorded in any source file, so these covariates could not be fully evaluated; and, because no missing values were imputed, the reduced sample directly limits statistical precision. Third, GF was used as a comparator rather than a validated gold-standard reference in goats, and visit-level quality information was unavailable; part of the observed LMD–GF disagreement may therefore originate on the GF side, and this study cannot apportion error between the two instruments. Fourth, LMD readings were four brief spot measurements whereas GF values represent a different sampling window, and calibration was empirical rather than physically grounded. Fifth, the ROC analysis was observation-level rather than animal-level, and the thresholds (13 g CH_4/day and a Youden-optimal value from a median split) lack operational validation, so the diagnostic results are exploratory and not a validated screening protocol. Sixth, data were collected at a single site over a 19-day window with several very windy days (up to 69.8 mph); the absence of a significant wind association ($n = 112$) does not show that wind has no effect on LMD accuracy [25,37]. Finally, the results should not be generalized to other goat breeds, production systems, or age classes without further validation.

Goat-specific LMD validation data remain limited [35], and GF has been evaluated in goats in only two studies from one institute [22,23]; the present dataset therefore provides an initial comparator analysis for this species. Larger multi-site, multi-season studies are needed, incorporating standardized LMD protocols (fixed sensor-to-muzzle distance, angle, and timing), physical LMD calibration, pre-specified thresholds, and animal-level phenotyping. Direct validation of both technologies against respiration chambers [7,8] would separate GF sampling error from LMD error; in such a design, GF should be operated according to goat-specific recommendations (at least 18 goats per unit, at least four scheduled daily measurement windows, and visit numbers recorded for every animal-day [22,23]). Including animal-level covariates, such as body weight, body condition score, and FAMACHA score, in calibration models may also help explain residual variance [56]. Low-cost, field-deployable methane phenotyping tools remain highly desirable for genetic selection, mitigation research, and farm-level implementation [4–6,55].

5. Conclusions

In this exploratory, retrospective comparison of 26 Spanish Boer goats, no detectable agreement was found between laser methane detector (LMD) and GreenFeed (GF) measurements. Correlation, regression, Bland–Altman, and classification analyses all indicated negligible association and near-chance discrimination. Under the conditions tested, the LMD cannot be recommended as a replacement or screening tool for methane phenotyping in goats. Because the sample was small and GF was used as a comparator rather than a validated reference, these findings should be interpreted cautiously and viewed as exploratory rather than definitive. Larger studies using standardized LMD protocols and independent reference methods, such as respiration chambers, are needed to determine whether useful calibration relationships can be established in goats.

Author Contributions: Conceptualization, B.R.M. and H.I.; methodology, B.R.M., N.G., and H.I.; software, H.I. and B.R.M.; formal analysis, B.R.M. and H.I.; investigation, M.H., S.C., V.K., H.A., B.B.; H.I. and B.R.M.; resources, B.R.M.; data curation, M.H., S.C., V.K., H.A., B.B.; writing—original draft preparation, H.I. and B.R.M.; writing—review and editing, B.R.M. and H.I.; visualization, H.I.

and B.R.M.; supervision, B.R.M.; validation, B.R.M. and H.I.; funding acquisition, B.R.M.; project administration, B.R.M. All authors have read and agreed to the published version of the manuscript.

Funding: This research was funded by the USDA/NRCS: NR233A750004G103.

Institutional Review Board Statement: The study was approved by the Tuskegee University Animal Care and Use Committee (TUACUC) (protocol code R08-2024-06) on 30 August 2024.

Informed Consent Statement: Not applicable.

Data Availability Statement: The data used for this study are available from the corresponding author upon request. The data that supports the findings of this research are not publicly available due to ongoing intellectual property protection considerations.

Acknowledgments: Thanks to the Cooperative Extension Program for Agricultural Environmental Sciences, Tuskegee University, Tuskegee, AL, USA.

Conflicts of Interest: The authors declare no conflicts of interest.

References

1. IPCC. Climate Change 2021: The Physical Science Basis. In *Contribution of Working Group I to the Sixth Assessment Report of the Intergovernmental Panel on Climate Change*; Cambridge University Press: Cambridge, UK, 2021.
2. IPCC. Climate Change 2014: Synthesis Report. In *Contribution of Working Groups I, II and III to the Fifth Assessment Report of the Intergovernmental Panel on Climate Change*; Pachauri, R.K., Meyer, L.A., Eds.; IPCC: Geneva, Switzerland, 2014.
3. Hammond, K.J.; Crompton, L.A.; Bannink, A.; Dijkstra, J.; Yáñez-Ruiz, D.R.; O'Kiely, P.; Kebreab, E.; Eugène, M.A.; Yu, Z.; Shingfield, K.J.; et al. Review of current in vivo measurement techniques for quantifying enteric methane emission from ruminants. *Anim. Feed Sci. Technol.* **2016**, *219*, 13–30. [CrossRef]
4. Lassen, J.; Difford, G.F. Review: Genetic and genomic selection as a methane mitigation strategy in dairy cattle. *Animal* **2020**, *14*, s473–s483. [CrossRef] [PubMed]
5. Difford, G.F.; Plichta, D.R.; Løvendahl, P.; Lassen, J.; Noel, S.J.; Højberg, O.; Wright, A.-D.G.; Zhu, Z.; Kristensen, L.; Nielsen, H.B.; et al. Host genetics and the rumen microbiome jointly associate with methane emissions in dairy cows. *PLoS Genet.* **2018**, *14*, e1007580. [CrossRef] [PubMed]
6. Reisinger, A.; Clark, H.; Cowie, A.L.; Emmet-Booth, J.; Gonzalez Fischer, C.; Herrero, M.; Howden, M.; Leahy, S. How necessary and feasible are reductions of methane emissions from livestock to support stringent temperature goals? *Philos. Trans. R. Soc. A* **2021**, *379*, 20200452. [CrossRef] [PubMed]
7. Johnson, K.A.; Johnson, D.E. Methane emissions from cattle. *J. Anim. Sci.* **1995**, *73*, 2483–2492. [CrossRef] [PubMed]
8. Storm, I.M.L.D.; Hellwing, A.L.F.; Nielsen, N.I.; Madsen, J. Methods for measuring and estimating methane emission from ruminants. *Animals* **2012**, *2*, 160–183. [CrossRef] [PubMed]
9. Grainger, C.; Clarke, T.; McGinn, S.M.; Auldist, M.J.; Beauchemin, K.A.; Hannah, M.C.; Waghorn, G.C.; Clark, H.; Eckard, R.J. methane emissions from dairy cows measured using the sulfur hexafluoride (SF_6) tracer and chamber techniques. *J. Dairy Sci.* **2007**, *90*, 2755–2766. [CrossRef] [PubMed]
10. Zhao, Y.; Aubry, A.; O'Connell, N.E.; Annett, R.; Yan, T. A review of enteric methane emission measurement techniques in ruminants. *Animals* **2020**, *10*, 1004. [CrossRef] [PubMed]
11. Patra, A.K. Recent advances in measurement and dietary mitigation of enteric methane emissions in ruminants. *Front. Vet. Sci.* **2016**, *3*, 39. [CrossRef] [PubMed]
12. Global Research Alliance on Agricultural Greenhouse Gases. *Guidelines for Estimating methane Emissions from Individual Ruminants Using: GF, 'Sniffers', Hand-Held Laser Detector and Portable Accumulation Chambers*; Jonker, A., Waghorn, G.C., Eds.; New Zealand Agricultural Greenhouse Gas Research Centre: Palmerston North, New Zealand, 2020.
13. Hill, J.; McSweeney, C.; Wright, A.-D.G.; Bishop-Hurley, G.; Kalantar-zadeh, K. Measuring methane production from ruminants. *Trends Biotechnol.* **2016**, *34*, 26–35. [CrossRef] [PubMed]
14. Negussie, E.; Mehtiö, T.; Mäntysaari, P.; Løvendahl, P.; Mäntysaari, E.A.; Lidauer, M.H. Non-invasive individual methane measurement in dairy cows. *Animal* **2017**, *11*, 890–899. [CrossRef] [PubMed]
15. Hammond, K.J.; Humphries, D.J.; Crompton, L.A.; Green, C.; Reynolds, C.K. Methane emissions from cattle: Estimates from short-term measurements using a GreenFeed system compared with respiration chambers or sulphur hexafluoride tracer. *Anim. Feed Sci. Technol.* **2015**, *203*, 41–52. [CrossRef]
16. Hristov, A.N.; Oh, J.; Giallongo, F.; Frederick, T.W.; Harper, M.T.; Weeks, H.L.; Branco, A.F.; Moate, P.J.; Deighton, M.H.; Williams, S.R.O.; et al. Comparison of the GreenFeed system with the sulfur hexafluoride tracer technique for measuring enteric methane emissions from dairy cows. *J. Dairy Sci.* **2016**, *99*, 5461–5465. [CrossRef] [PubMed]

17. Arbre, M.; Rochette, Y.; Guyader, J.; Lascoux, C.; Gómez, L.M.; Eugène, M.; Morgavi, D.P.; Renand, G.; Doreau, M.; Martin, C. Repeatability of enteric methane determinations from cattle using either the SF_6 tracer technique or the GreenFeed system. *Anim. Prod. Sci.* **2016**, *56*, 238–243. [CrossRef]
18. Hristov, A.N.; Kebreab, E.; Niu, M.; Oh, J.; Bannink, A.; Bayat, A.R.; Boland, T.M.; Brito, A.F.; Casper, D.P.; Crompton, L.A.; et al. Symposium review: Uncertainties in enteric methane inventories, measurement techniques, and prediction models. *J. Dairy Sci.* **2018**, *101*, 6655–6674. [CrossRef] [PubMed]
19. Huhtanen, P.; Ramin, M.; Hristov, A.N. Enteric methane emission can be reliably measured by the GreenFeed monitoring unit. *Livest. Sci.* **2019**, *222*, 31–40. [CrossRef]
20. Huhtanen, P.; Cabezas-Garcia, E.H.; Utsumi, S.; Zimmerman, S. Comparison of methods to determine methane emissions from dairy cows in farm conditions. *J. Dairy Sci.* **2015**, *98*, 3394–3409. [CrossRef] [PubMed]
21. Alemu, A.W.; Vyas, D.; Manafiazar, G.; Basarab, J.A.; Beauchemin, K.A. Enteric methane emissions from low- and high-residual feed intake beef heifers measured using GreenFeed and respiration chamber techniques. *J. Anim. Sci.* **2017**, *95*, 3727–3737. [CrossRef]
22. Tadesse, D.; Puchala, R.; Yirga, H.; Patra, A.K.; Gipson, T.A.; Min, B.R.; Goetsch, A.L. Determining appropriate numbers and times of daily measurements using GreenFeed system to estimate ruminal methane emission of meat goats. *Animals* **2024**, *14*, 835. [CrossRef] [PubMed]
23. Tadesse, D.; Puchala, R.; Yirga, H.; Patra, A.K.; Goetsch, A.L. Effect of group size in a Greenfeed system on estimates of enteric methane emissions from goats grazing grass-based pasture. *Smart Agric. Technol.* **2025**, *11*, 100989. [CrossRef]
24. Iseki, T. A portable remote methane detector using an InGaAsP DFB laser. *Environ. Geol.* **2004**, *46*, 1064–1069. [CrossRef]
25. Chagunda, M.G.G.; Ross, D.; Roberts, D.J. On the use of a laser methane detector in dairy cows. *Comput. Electron. Agric.* **2009**, *68*, 157–160. [CrossRef]
26. Chagunda, M.G.G.; Yan, T. Do methane measurements from a laser detector and an indirect open-circuit respiration calorimetric chamber agree sufficiently closely? *Anim. Feed Sci. Technol.* **2011**, *165*, 8–14. [CrossRef]
27. Ricci, P.; Chagunda, M.G.G.; Rooke, J.; Houdijk, J.G.M.; Duthie, C.-A.; Hyslop, J.; Roehe, R.; Waterhouse, A. Evaluation of the laser methane detector to estimate methane emissions from ewes and steers. *J. Anim. Sci.* **2014**, *92*, 5239–5250. [CrossRef] [PubMed]
28. Grobler, S.M.; Scholtz, M.M.; van Rooyen, H.; Mpayipheli, M.; Neser, F.W.C. Methane production in different breeds, grazing different pastures or fed a total mixed ration, as measured by a laser methane detector. *S. Afr. J. Anim. Sci.* **2014**, *44*, 12–16. [CrossRef]
29. Pickering, N.K.; Chagunda, M.G.G.; Banos, G.; Mrode, R.; McEwan, J.C.; Wall, E. Genetic parameters for predicted methane production and laser methane detector measurements. *J. Anim. Sci.* **2015**, *93*, 11–20. [CrossRef] [PubMed]
30. Mapfumo, L.; Grobler, S.; Mupangwa, J.; Scholtz, M.; Muchenje, V. Enteric methane output from selected herds of beef cattle raised under extensive arid rangelands. *Pastoralism* **2018**, *8*, 15. [CrossRef]
31. Wu, L. Measurement Methods to Assess Methane Production of Individual Dairy Cows in a Barn. Ph.D. Thesis, Wageningen University and Research, Wageningen, The Netherlands, 2016.
32. Robinson, D.L.; Goopy, J.P.; Hegarty, R.S.; Oddy, V.H. Comparison of repeated measurements of methane production in sheep over 5 years and a range of measurement protocols. *J. Anim. Sci.* **2015**, *93*, 4637–4650. [CrossRef] [PubMed]
33. Reintke, J.; Brügemann, K.; Yin, T.; Engel, P.; Wagner, H.; Wehrend, A.; König, S. Assessment of methane emission traits in ewes using a laser methane detector: Genetic parameters and impact on lamb weaning performance. *Arch. Anim. Breed.* **2020**, *63*, 113–123. [CrossRef] [PubMed]
34. Doran, S. Development of Proxy Indicators of methane Output by Sheep Using Rapid-Throughput Field and Laboratory Technologies. Ph.D. Thesis, Aberystwyth University, Aberystwyth, UK, 2014.
35. Roessler, R.; Chefor, F.; Schlecht, E. Using a portable laser methane detector in goats to assess diurnal, diet- and position-dependent variations in enteric methane emissions. *Comput. Electron. Agric.* **2018**, *150*, 110–117. [CrossRef]
36. Sorg, D.; Difford, G.F.; Mühlbach, S.; Kuhla, B.; Swalve, H.H.; Lassen, J.; Strabel, T.; Pszczola, M. Comparison of a laser methane detector with the GreenFeed and two breath analysers for on-farm measurements of methane emissions from dairy cows. *Comput. Electron. Agric.* **2018**, *153*, 285–294. [CrossRef]
37. Chagunda, M.G.G.; Flockhart, J.F.; Roberts, D.J. Measurement of enteric methane from ruminants using a hand-held laser methane detector. *Acta Agric. Scand. A Anim. Sci.* **2013**, *63*, 68–75. [CrossRef]
38. Sorg, D.; Mühlbach, S.; Rosner, F.; Kuhla, B.; Derno, M.; Meese, S.; Schwarm, A.; Kreuzer, M.; Swalve, H. The agreement between two next-generation laser methane detectors and respiration chamber facilities in recording methane concentrations in the spent air produced by dairy cows. *Comput. Electron. Agric.* **2017**, *143*, 262–272. [CrossRef]
39. Beauchemin, K.A.; Ungerfeld, E.M.; Eckard, R.J.; Wang, M. Review: Fifty years of research on rumen methanogenesis: Lessons learned and future challenges for mitigation. *Animal* **2020**, *14*, s2–s16. [CrossRef] [PubMed]
40. Bland, J.M.; Altman, D.G. Statistical methods for assessing agreement between two methods of clinical measurement. *Lancet* **1986**, *327*, 307–310. [CrossRef]

41. Bland, J.M.; Altman, D.G. Measuring agreement in method comparison studies. *Stat. Methods Med. Res.* **1999**, *8*, 135–160. [CrossRef] [PubMed]
42. Bland, J.M.; Altman, D.G. Agreement between methods of measurement with multiple observations per individual. *J. Biopharm. Stat.* **2007**, *17*, 571–582. [CrossRef] [PubMed]
43. Myles, P.S.; Cui, J. Using the Bland–Altman method to measure agreement with repeated measures. *Br. J. Anaesth.* **2007**, *99*, 309–311. [CrossRef] [PubMed]
44. Choudhary, P.K.; Nagaraja, H.N. *Measuring Agreement: Models, Methods, and Applications*; John Wiley & Sons: Hoboken, NJ, USA, 2017.
45. Goopy, J.P.; Robinson, D.L.; Woodgate, R.T.; Donaldson, A.J.; Oddy, V.H.; Vercoe, P.E.; Hegarty, R.S. Estimates of repeatability and heritability of methane production in sheep using portable accumulation chambers. *Anim. Prod. Sci.* **2015**, *56*, 116–122. [CrossRef]
46. Percie du Sert, N.; Hurst, V.; Ahluwalia, A.; Alam, S.; Avey, M.T.; Baker, M.; Browne, W.J.; Clark, A.; Cuthill, I.C.; Dirnagl, U.; et al. The ARRIVE guidelines 2.0: Updated guidelines for reporting animal research. *PLoS Biol.* **2020**, *18*, e3000410. [CrossRef] [PubMed]
47. Kilkenny, C.; Browne, W.J.; Cuthill, I.C.; Emerson, M.; Altman, D.G. Improving bioscience research reporting: The ARRIVE guidelines for reporting animal research. *PLoS Biol.* **2010**, *8*, e1000412. [CrossRef] [PubMed]
48. Cohen, J. A coefficient of agreement for nominal scales. *Educ. Psychol. Meas.* **1960**, *20*, 37–46. [CrossRef]
49. Landis, J.R.; Koch, G.G. The measurement of observer agreement for categorical data. *Biometrics* **1977**, *33*, 159–174. [CrossRef]
50. McNemar, Q. Note on the sampling error of the difference between correlated proportions or percentages. *Psychometrika* **1947**, *12*, 153–157. [CrossRef] [PubMed]
51. Hanley, J.A.; McNeil, B.J. The meaning and use of the area under a receiver operating characteristic (ROC) curve. *Radiology* **1982**, *143*, 29–36. [CrossRef] [PubMed]
52. Bates, D.; Mächler, M.; Bolker, B.; Walker, S. Fitting linear mixed-effects models using lme4. *J. Stat. Softw.* **2015**, *67*, 1–48.
53. Youden, W.J. Index for rating diagnostic tests. *Cancer* **1950**, *3*, 32–35. [CrossRef]
54. Goopy, J.P.; Woodgate, R.; Donaldson, A.; Robinson, D.L.; Hegarty, R.S. Validation of a short-term methane measurement using portable static chambers to estimate daily methane production in sheep. *Anim. Feed Sci. Technol.* **2011**, *166–167*, 219–226. [CrossRef]
55. Min, B.R.; Solaiman, S.; Waldrip, H.M.; Parker, D.; Todd, R.W.; Brauer, D. Dietary mitigation of enteric methane emissions from ruminants: A review of plant tannin mitigation options. *Anim. Nutr.* **2020**, *6*, 231–246. [CrossRef] [PubMed]
56. Min, B.-R.; Lee, S.; Jung, H.; Miller, D.N.; Chen, R. Enteric Methane Emissions and Animal Performance in Dairy and Beef Cattle Production: Strategies, Opportunities, and Impact of Reducing Emissions. *Animals* **2022**, *12*, 948. [CrossRef] [PubMed]

Disclaimer/Publisher's Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.