



Reliable biochemical methane potential testing: insights and recommendations from global interlaboratory study

Sasha D. Hafner^{a,*}, Estelle Leca^{b,c}, Cecilia Sambusiti^d, Sergi Astals^e, Olivier Azam^f, Konrad Koch^g, Jing Liu^h, Henrik B. Møller^a, Mihaela Nistorⁱ, Mario Olaya-Rincon^e, Christine Peyrelasseⁱ, Alastair J. Ward^a, Bastien Zennaroⁱ, Florian Monlau^b

^a Department of Biological and Chemical Engineering, Aarhus University, Aarhus, Denmark

^b TotalEnergies, OneTech R&D, PERL ESD – Pôle D'Etudes et de Recherche de Lacq, Pôle Economique 2, BP 47 – RD 817, 64170 Lacq, France

^c AKKODIS, Technopôle HELIOPARC Bat Newton, 4 rue Jules FERRY, 64000 Pau, France

^d TotalEnergies, Onetech R&D, CSTJF – Avenue Larribau, 64018 Pau, France

^e Department of Chemical Engineering and Analytical Chemistry, Universitat de Barcelona 08028 Barcelona, Spain

^f INRAE Transfert METYS, 60 rue Nicolas Leblanc, 11 100 Narbonne, France

^g Chair of Urban Water Systems Engineering, Technical University of Munich, Am Coulombwall 3, 85748 Garching, Germany

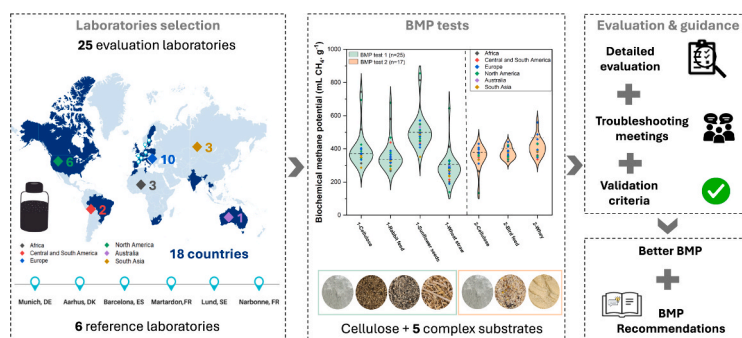
^h BPC Instruments AB, Mobilvägen 10, 223 62 Lund, Sweden

ⁱ APESA, 3 Chemin de Sers, 64121 Montardon, France

HIGHLIGHTS

- 31 laboratories from 17 countries measured biochemical methane potential in 2 tests.
- Reproducibility standard deviation was <12% for 6 experienced reference laboratories.
- Test 1 reproducibility was poor for other laboratories (21–33%).
- Performance improved after evaluation, troubleshooting, and method modifications.
- Validation criteria and SMP curve screening provided additional improvement.

GRAPHICAL ABSTRACT



ABSTRACT

Thirty-one laboratories measured biochemical methane potential (BMP) of six substrates in a project aimed at improving BMP measurement quality. Laboratories used their established measurement procedures, with limited standardization. As in earlier studies, reproducibility was quantified, but here laboratories came from a wide geographic range, and the accuracy of BMP measurements was assessed by comparison with 6 experienced “reference” laboratories. Measurement reproducibility between the remaining “evaluation” laboratories was poor in the first test; relative reproducibility standard deviation was 21–33% for four substrates. Values improved to 9–19% in the second test following review of a detailed analysis of submitted data, troubleshooting meetings, and method modifications by some laboratories. Reference laboratories performed better, with reproducibility standard deviation of 5–12%. Repeatability within laboratories was also better for the reference group, with relative standard deviation of 2–4% vs. 3–7% for evaluation laboratories. Evaluation laboratory results showed a persistent negative bias of 4–16% compared to reference laboratories. Application of cellulose validation criteria improved reproducibility but not accuracy. Most of the extreme BMP values were associated with abnormal methane production curves, and a subjective visual evaluation proved valuable as a screening tool. The cause of measurement error

* Corresponding author.

E-mail address: sasha.hafner@bce.au.dk (S.D. Hafner).

<https://doi.org/10.1016/j.biortech.2026.134136>

Received 24 September 2025; Received in revised form 29 January 2026; Accepted 29 January 2026

Available online 1 February 2026

0960-8524/© 2026 The Author(s). Published by Elsevier Ltd. This is an open access article under the CC BY license (<http://creativecommons.org/licenses/by/4.0/>).

was generally unknown, although data processing and volatile solids measurement errors caused some extreme results, and changes in inoculum source improved performance in two cases. There were no consistent differences among the diverse measurement methods and the sources of error probably vary among laboratories. Consistently accurate BMP measurement is possible for laboratories that follow established guidelines and develop experience.

1. Introduction

Biochemical methane potential (BMP) is commonly used to assess substrate suitability for biogas production and is used in both research and the biogas industry. Diverse methods are employed for measuring biogas quantity in BMP tests, including methods based on measurement of volume, pressure, or mass. Because only part of the biogas is methane (CH_4), removal of carbon dioxide (CO_2) or determination of the biogas CH_4 content is necessary, and gas chromatography and infrared gas analyzers are common for manual methods, while CO_2 capture in an alkaline solution is often used with automated systems.

Reported values for interlaboratory reproducibility standard deviation in BMP (among laboratories) vary: 15–37% (Raposo et al., 2011), 6–19% (Weinrich et al., 2018; Paterson et al., 2020), 18% (Ribeiro et al., 2020), and 8–24% (Hafner et al., 2020). These better values (10% or lower) are precise considering the complexity of the analysis. However, a small number of extreme BMP measurements are common (and sometimes excluded from analysis); previous studies showed a factor of two or greater between the highest and lowest measurements (Raposo et al., 2011; Weinrich et al., 2018; Ribeiro et al., 2020; Hafner et al., 2020).

Standardization of measurement methods is one approach to improving accuracy. The German VDI guidelines (VDI, 2006, 2016) include detailed recommendations for several measurement methods along with many general requirements, including an inoculum-to-substrate ratio of at least 2:1 (volatile solids (VS) basis). The VDLUFA (2011) guidelines are similar but simpler and the set from Holliger et al. (2016, 2021a) even more so. These and other guidelines are based on a combination of empirical evidence from research projects and expert opinions. The VDLUFA (2011) method introduced the use of microcrystalline cellulose as a “reference substrate” or “positive control”, with limits of 670.5 and 819.5 mL g^{-1} as biogas per g substrate VS, based on the expected theoretical value $\pm 10\%$ (335.25–409.75 mL g^{-1} as methane, assuming 50% methane in biogas as in VDI (2016)). The later VDI (2016) and Holliger et al. (2016, 2021a) guidelines follow suit but explicitly state that microcrystalline cellulose should be included in each test or batch. A large interlaboratory study confirmed that discarding tests that did not meet cellulose BMP criteria eliminated extreme BMP values for other complex substrates and improved reproducibility (Hafner et al., 2020). Although the use of a cellulose positive control can help, without comparison to other laboratories, persistent bias for complex substrates is possible. Participation in interlaboratory comparisons such as the annual German KTBL/VDLUFA proficiency test (<https://www.ktbl.de/themen/proficiencytest>) can help in this regard (Paterson et al., 2020).

This present work continued the effort to understand and improve reproducibility and accuracy in BMP measurement, with some differences from earlier studies. Previous interlaboratory projects on BMP have been restricted to Europe or nearly so (Raposo et al., 2011; Weinrich et al., 2018; Ribeiro et al., 2020; Hafner et al., 2020; Paterson et al., 2020). But biogas technology is widely used at household and industrial scales worldwide (Garfí et al., 2016; Kemausuor et al., 2018; Xue et al., 2020; Seminario-Córdova, 2023). This present work included laboratories from a broader geographic area, including the Americas, Asia, Africa, and Oceania, and should better reflect the global quality of BMP testing. A second distinctive feature was the effort to improve the BMP measurement accuracy of participating laboratories through detailed evaluation of submitted BMP test data and troubleshooting between two sets of BMP tests. This component provided a means to test the hypothesis that laboratory personnel can improve measurement

accuracy based on an external evaluation. Troubleshooting and method standardization are not new for BMP (Paterson et al., 2020; Ribeiro et al., 2020; Hafner et al., 2020), but the present work included an evaluation of original raw measurement data and not only final BMP values. Furthermore, by inviting laboratories to participate, the selection bias resulting from laboratories themselves initiating participation, as in the KTBL/VDLUFA proficiency program, was likely reduced. The third distinction from previous work was a more direct assessment of measurement accuracy and not only reproducibility. Participating laboratories were divided into two groups at the start, with 6 experienced laboratories providing “reference” BMP results based on previous experience, and evaluation of BMP reproducibility and accuracy focused on the remaining 25 laboratories. This structure also provided a way to test the hypothesis that experience is a predictor of measurement reproducibility.

The objectives of this study were to (1) quantify reproducibility and accuracy in BMP measurement among a global network of laboratories, (2) identify and quantify factors that affect BMP accuracy and reproducibility, and (3) test whether detailed evaluation and troubleshooting can improve the quality of BMP measurement.

2. Methods

Participating laboratories carried out one or two BMP tests with multiple substrates, with feedback and an opportunity for troubleshooting with project organizers between the two tests. Laboratories submitted underlying raw measurement data in addition to BMP values. Measurement reproducibility was assessed directly by calculation of reproducibility standard deviation between laboratory mean BMP values, and the accuracy of BMP measured by “evaluation” laboratories was assessed by comparison with results from “reference” laboratories. Detailed measurement data were used to evaluate the causes of variability.

2.1. Evaluation and reference laboratories

Participating laboratories included academic, commercial, and government research institutions located in Australia, Belgium, Brazil, Canada, Denmark, Germany, Spain, Sweden, France, Italy, Morocco, Poland, South Africa, Thailand, Tunisia, the United Kingdom, and the United States. Prospective laboratories were identified through professional networks and research papers. In total, 25 “evaluation” laboratories participated. Six “reference” laboratories were selected based on experience. Reference laboratories had participated in previous European interlaboratory studies or related comparisons (Raposo et al., 2011; Ribeiro et al., 2020; Hafner et al., 2020; Justesen et al., 2019; Hafner and Astals, 2019) and had extensive experience in BMP measurement. Author affiliations a, e, f, g, b, and i were reference laboratories. Evaluation laboratories were promised confidentiality and therefore additional laboratory details are not provided here.

2.2. BMP tests and method improvements

Two rounds of BMP tests were completed, with all reference laboratories and most evaluation laboratories participating in both. Test 2 provided an opportunity for the evaluation laboratories to demonstrate an improvement in their methods, based on a detailed analysis of their test 1 results and an opportunity to discuss methods and possible improvements with project organizers. Feedback on both tests was provided in the form of a detailed report.

BMP test 1 included three complex substrates and microcrystalline cellulose (Sigma Aldrich, Burlington MA USA) as a positive control, while test 2 included two complex substrates and the same microcrystalline cellulose (Table 1). Substrates were selected to provide different BMP values and variation in chemical composition to simulate different types of practical substrates. Except for cellulose, the composition of substrates was not disclosed to evaluation laboratories, and the use of different complex substrates for the two tests was done to avoid any bias in measurement or data processing due to knowledge of expected BMP values, following standard proficiency testing practice. To avoid practical problems with international shipping and changes in BMP during shipping or storage, all substrates were dry and ground (2 mm screen). Two different animal feeds were used: one for rabbits in test 1 ("Rabbit Pellets", GASCO, France) and one for birds in test 2 ("Four-Season Pigeon Grain Mix", GASCO, France). Wheat straw ("Natural Straw Bedding", GASCO, France) was included in test 1 as a slowly degradable substrate. Sunflower seeds ("Striped Sunflower Seeds", GASCO, France) in test 1 were high in lipids, while whey powder ("Neutral-Flavored Whey Protein Powder", CORENGTH, France) in test 2 was the only substrate primarily composed of protein. While this design provided an opportunity to assess laboratory performance with a wide variety of substrates and ensured that participants would not know the expected BMP of any of the complex substrates, having different complex substrates in the two tests complicated assessment of improvement.

The contents of carbon (C), hydrogen (H), nitrogen (N), and sulfur (S) in all complex substrates were measured using a Vario Microcube (Elementar, Germany) by project organizers and oxygen (O) was calculated by difference. Elemental composition was used to estimate the theoretical maximum BMP, assuming complete degradation and no partitioning of substrate to microbial biomass, with the predBg() function from the biogas package in R (Hafner et al., 2018), which is based on Rittmann and McCarty (2020: Eq. 10.5). Total protein was estimated as 6.25 times the N content. Lipids were measured by Soxhlet extraction with petroleum ether as the solvent. Klason Lignin was analyzed using acid hydrolysis following the NREL protocol with HPLC for quantification (Sluiter et al., 2008). Total carbohydrates were estimated by difference.

A minimal BMP protocol was shared with all participating laboratories to standardize BMP methods to a limited degree. Laboratories were asked to: (1) measure total solids (TS) and VS of inoculum and substrate in triplicate, (2) measured pH of the inoculum, (3) use a VS-based inoculum-to-substrate ratio of 2:1, (4) include microcrystalline cellulose as a positive control, (5) use triplicates for BMP testing, (6) include "blanks" with only inoculum for calculating BMP, (7) incubate at a constant temperature between 35 and 40 °C that matched the temperature of the inoculum source, (8) run the test to the 1% net rate criterion following Holliger et al. (2021b), and (9) submit calculated BMP, along with raw setup data and methane production data in the

appropriate template provided by project organizers. Each laboratory used its own inoculum but was encouraged to use an inoculum from an active anaerobic digester at steady state (Holliger et al., 2016). This protocol was minimal, recognizing that different approaches can provide accurate BMP results and aiming for a balance between laboratory staff effort on the one side and a combination of standardization for quality control and measurement of indicators that may provide insight into causes of variability in BMP measurement on the other. These requirements were roughly consistent with other guidelines (VDI, 2006; 2016; VDLUFA, 2011; Holliger et al., 2021a, b) but less comprehensive. For example, there was no requirement to reduce endogenous methane production from inocula by an incubation prior to starting the test ("degassing" or "preincubation") in contrast with VDLUFA (2011), because measurements show that effects of this step are generally small for typical inocula (Elbeshbishy et al., 2012; Hülsemann et al., 2020; Oliveira et al., 2022). This protocol did not include any validation criteria; instead, different criteria were assessed as part of data analysis by project organizers. Measurements not uniformly recommended across BMP guidelines, such as final pH and apparent VS destruction, were not included. Furthermore, it is likely that many laboratories, especially those in the reference group, followed more comprehensive guidance such as VDI (2016).

Laboratories used a range of methods to measure methane production. Among evaluation laboratories, manual methods dominated, with 10 laboratories using manual volumetric and 4 manometric methods. Eight laboratories used an automatic system (AMPTS®, BPC, Lund, Sweden). A small number used other automated systems (custom or commercial) or else a manual gravimetric method. Reference laboratories were similar, with only 2 using AMPTS® compared to 2 manual volumetric and 2 manual manometric. Two evaluation laboratories switched from manual volumetric or manometric for test 1 to the gravimetric method for test 2 based on feedback following evaluation of test 1 results. Apart from CO₂ removal used for AMPTS®, gas chromatography was used by 8 evaluation and 3 reference laboratories. Gas density (Justesen et al., 2019) was used by 1 reference laboratory.

All data were submitted using spreadsheet templates provided by organizers. As is typical in inter-laboratory comparisons, each participating laboratory submitted the BMP values that they determined, but atypically, also provided detailed raw measurement data that allowed project organizers to reproduce their calculations. The BMP values determined by each laboratory were referred to as "submitted" values, and were compared to "calculated" BMP values that project organizers calculated from raw data, as described in more detail in section 2.3, to identify potential calculation mistakes.

Following the first test, an online meeting was held with project organizers and reference laboratory staff, where preliminary results were presented and suggestions for evaluation laboratory troubleshooting and methodology improvements were developed. Each

Table 1

Substrate characteristics. Analytical methods are described in the text. Values for total solids (TS), volatile solids (VS), and biochemical methane potential (BMP) are median values from reference laboratory measurements. Theoretical maximum BMP was based on elemental composition, as given in the chemical formula, which is theoretical for cellulose and empirical for all others (based on elemental analysis). Apparent degradability was calculated as reference BMP divided by the theoretical maximum. BMP is expressed as dry methane volume at 101.325 kPa and 0 °C divided by substrate VS mass. Cellulose macromolecular composition (100% carbohydrates) was assumed, not measured. FM = fresh mass.

Test	Substrate	Chemical formula	TS (% FM)	VS (% TS)	Carbohydrates (% VS)	Proteins (% VS)	Lipids (% VS)	Lignin (% VS)	Theoretical maximum BMP (mL g ⁻¹)	Reference BMP (mL g ⁻¹)	Apparent degradability (%)
1	Cellulose	C ₆ H ₁₀ O ₅	99.0	100	100	0.0	0.0	0.0	414	375	91
1	Rabbit feed	C _{20.0} H _{38.0} O _{13.1} N	90.4	89.6	76.6	14.0	3.5	6.0	494	352	71
1	Wheat straw	C _{24.4} H _{43.4} O _{17.1} N	91.3	87.7	68.5	13.0	0.5	18.1	464	307	66
1	Sunflower seeds	C _{21.5} H _{39.5} O _{7.59} N	93.8	97.3	33.6	19.1	41.2	6.2	692	523	76
2	Cellulose	C ₆ H ₁₀ O ₅	99.2	100	100	0.0	0.0	0.0	414	377	91
2	Bird feed	C _{24.3} H _{50.4} O _{19.8} N	89.0	98.0	81.9	12.5	3.6	2.1	437	405	93
2	Whey	C _{4.80} H _{8.88} O _{1.95} N	94.5	96.6	22.3	72.4	5.4	0.0	530	465	88

evaluation laboratory was then sent a report with a detailed evaluation of their results and recommendations for improvements. At this point, evaluation laboratories were invited to discuss the issues encountered and potential solutions during the analysis through an online troubleshooting meeting. Not all evaluation laboratories accepted, and a small number submitted inaccurate results for both tests.

2.3. Data processing and evaluation

All data processing was performed in R (R Core Team, 2025). The publicly available biogas package v1.64 (<https://cran.r-project.org/package=biogas>) was used to calculate cumulative standardized methane volume from submitted measurement data using the functions `calcBgVol()` (volumetric), `calcBgMan()` (manometric), `calcBgGrav()` (gravimetric), and `calcBgGD()` (gas density) (Hafner et al., 2018). The `summBg()` function was used for calculation of specific methane production (SMP) and BMP as well as assessment of the 1% net rate criterion. Calculations followed the steps given in the Standard BMP Methods documents 200 through 204 (<https://www.dbfz.de/en/projects/bmp>). The raw data used in these calculations included “setup” data (the mass of inoculum and substrate in each bottle) and biogas quantity (e.g., bottle pressure, removed volume, or bottle mass change) and composition (methane content of biogas, or, for the gas density method, the biogas volume and mass). BMP was calculated from raw data in two different ways: using the reference VS values and using individual participating laboratory measurements of VS. BMP calculated with reference VS was used to eliminate the contribution of VS measurement error and values calculated with individual laboratory measurements were used to check the data processing calculations for each laboratory.

Reference values for TS and VS were taken as the median of the reference laboratory values. BMP was expressed as net dry methane volume at 101.325 kPa (1.0 atm) and 0 °C divided by added substrate VS mass throughout this work.

After data processing, three types of BMP values were available: (1) “submitted BMP”, (2) “calculated BMP, measured VS”, and (3) “calculated BMP, reference VS”. Inter-laboratory reproducibility was quantified as the standard deviation of laboratory-level mean values for BMP, TS, and VS. Relative standard deviation was calculated by dividing the standard deviation by the group mean. Accuracy was assessed by comparison to reference values, which were taken as the median values from the reference laboratories. Relative overall accuracy was quantified as bias, using the median of relative error, calculated as the difference between each laboratory mean result and the reference result divided by the reference result. For quantification of reproducibility and accuracy, submitted BMP was used.

Hypothesis testing for reproducibility was done using a bootstrap approach (Efron and Tibshirani, 1994), which accommodates the obvious non-normal error distribution (Fig. 1). A total of 10^4 resampling iterations were used, with separate median (for accuracy) and standard deviation (for reproducibility) values calculated from resampled reported BMP values for each laboratory type $\times$ test group. Estimates of p -values for comparisons were made using quantiles. Comparison of test 1 and 2 reproducibility results was complicated by the use of different complex substrates (Section 2.2). But cellulose was identical and the composition of the two animal feeds was similar (Table 1), so only these were used in two separate bootstrap analyses for reproducibility.

The two sets of calculated BMP values were used primarily to understand sources of error in BMP; differences between submitted BMP

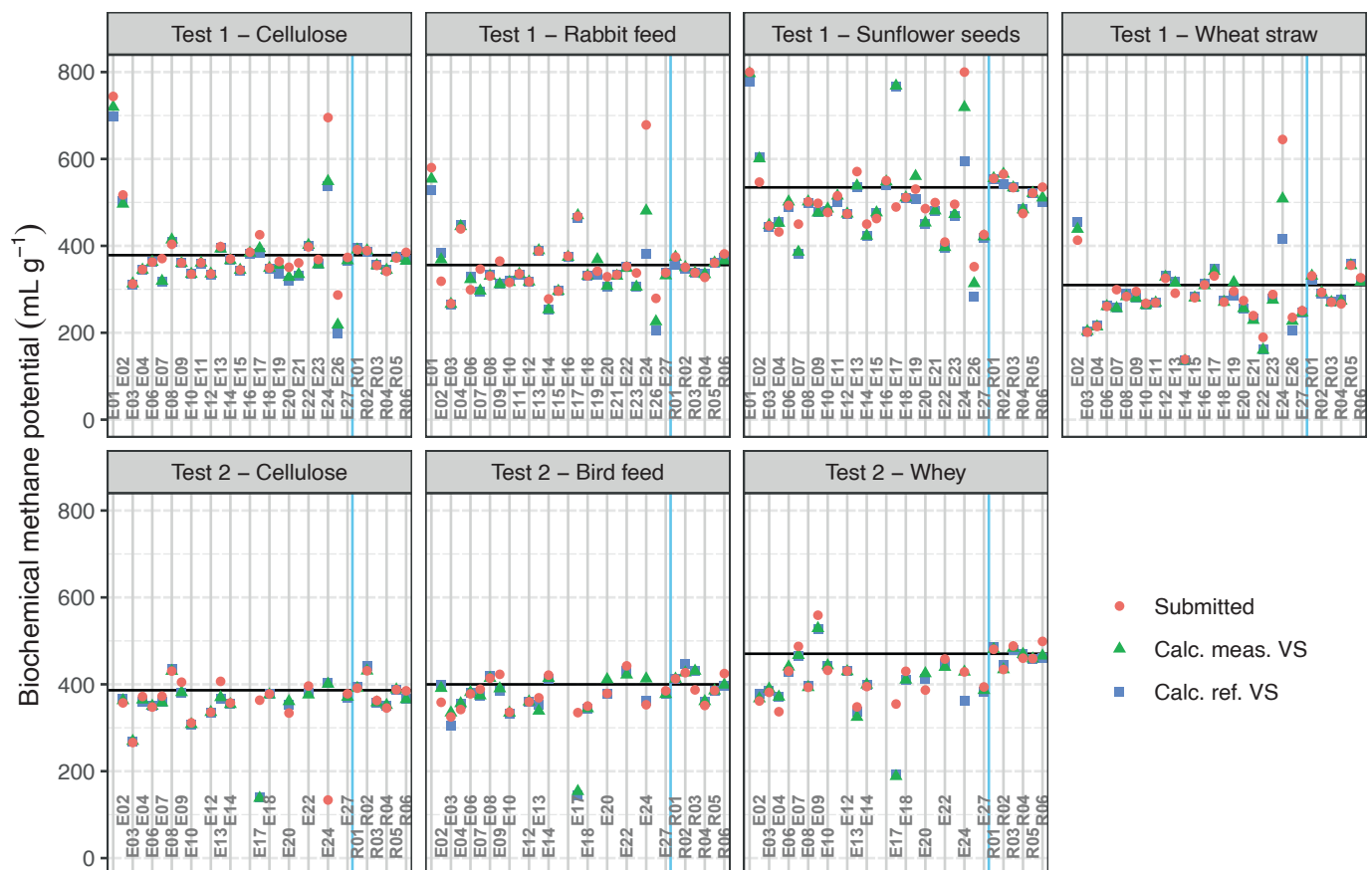


Fig. 1. Submitted BMP values and values calculated from raw data using either VS values measured by individual laboratories (“Calc. meas. VS.”) or reference VS values (“Calc. ref. VS”). Points to the right of the light blue line are reference laboratories. Horizontal (x axis) position is fixed by an anonymous laboratory ID key, with a prefix of E for evaluation laboratories and R for reference. Note that not all evaluation laboratories participated in Test 2. BMP is expressed as dry methane volume at 101.325 kPa and 0 °C divided by substrate VS mass. The heavy horizontal lines show reference BMP values.

and values calculated using measured VS were due only to calculation or data entry errors, while any additional difference between submitted BMP and values calculated using reference VS was due to VS measurement error.

Two sets of validation criteria were assessed by applying them to quantify their efficacy in improving reproducibility and accuracy: (1) those proposed by Holliger et al. (2021a), referred to as the “2021 criteria”, and (2) a set with less stringent criteria referred to as “relaxed criteria”. The 2021 set includes several criteria that were requested for all submitted data (see Table 1 in Holliger et al., 2021a): replication, use of cellulose, BMP units, minimum duration, and calculations. These two sets of criteria applied only differed in limits for mean cellulose BMP and the associated relative standard deviation among replicates. Cellulose BMP limits in the 2021 criteria are 340 and 395 mL g⁻¹, and maximum relative standard deviation among replicates 6%. The relaxed criteria used 330 and 415 mL g⁻¹, which is close to the reference BMP measured in this work $\pm 10\%$ as well as the apparent VDI (2016) and VDLUFA (2011) criteria as methane, and a 15% limit on relative standard deviation. Both criteria used here included a minimum duration when net methane production rate dropped below 1% of the cumulative value per day.

Relationships between SMP curve characteristics and BMP were explored using linear models with the lm() function in R. Kinetic results were extracted from each SMP curve using the fitFOM() first-order model function in the biogas package (Hafner et al., 2018) and included lag phase duration (as defined in Hafner et al., 2022) and a first-order rate constant. Nash-Sutcliffe model efficiency (NSE) was used to quantify the fit of each SMP curve to a first-order response (Nash and Sutcliffe, 1970). For blanks (inoculum-only bottles) the average specific methane production rate over the first three days was also determined, based on interpolated cumulative methane production. Calculations were carried out at the bottle level. Qualitative evaluation of SMP curves was also carried out by graphical evaluation of curves from all 619 bottles with substrate. Curves were first normalized by maximum SMP (so the maximum value was unity) to eliminate any impact of magnitude on the necessarily subjective rating. Curves were flagged as problematic if they showed a shape obviously different from a typical first-order response or a high rate of methane production at the end of the test. Short lag phases (up to 3 d) and a slight decline in SMP over time were accepted, but negative SMP was not. The evaluation was subjective, and several borderline curves could be identified in results.

3. Results and discussion

3.1. BMP reproducibility and accuracy

Submitted BMP values varied widely in test 1; differences between extreme values exceeded a factor of 2 for all substrates (Fig. 1). Reproducibility standard deviation was 21–33% for the evaluation group depending on the substrate (Table 2). Reference laboratories showed much better reproducibility: 5–12% (Table 2, Fig. 3). But the gap between evaluation and reference laboratories decreased from test 1 to test 2, with evaluation laboratory reproducibility standard deviation dropping to 9–19% in test 2 (Table 2, Fig. 3). The bootstrap analysis confirmed better reproducibility by reference laboratories in test 1 ($p = 0.001$ for cellulose, $p < 0.001$ for animal feed) and only for cellulose in test 2 ($p = 0.049$, $p = 0.18$ for animal feed). The analysis also confirmed an improvement in evaluation laboratory reproducibility for the animal feed substrates from test 1 to 2 ($p = 0.01$) but the evidence for the simpler cellulose substrate was inconclusive ($p = 0.054$). Comparing only the results for laboratories that participated in both tests show similar patterns: 23 vs. 19% for cellulose and 27 vs. 9% for the animal feeds.

The overall reproducibility improvement was driven by a moderate number of laboratories; of those that participated in both tests, the six laboratories with the highest average error across all substrates in test 1

Table 2

Interlaboratory reproducibility and accuracy in biochemical methane potential (BMP) for both tests by substrate and laboratory status (reference or evaluation). Reproducibility was quantified as the relative standard deviation in BMP values between laboratories, and error calculated by a comparison to reference values.

Test	Substrate	Evaluation laboratory n	Evaluation laboratory error (% of ref. BMP)	Reproducibility relative standard deviation (% of group mean)	
				Evaluation laboratories	Reference laboratories (n = 6)
1	Cellulose	25	−3.8	27	5.5
1	Rabbit feed	25	−5.7	26	5.9
1	Sunflower seeds	25	−7.8	21	6.0
1	Wheat straw	24	−11	33	12
2	Cellulose	17	−5.9	19	7.6
2	Bird feed	17	−7.7	9.3	7.3
2	Whey	17	−16	13	5.0

(laboratories E02, E03, E04, E14, E22, and E24) showed a reduction in average error going to test 2.

On average, evaluation laboratories tended to have a small negative bias in BMP results (based on comparison to reference values). Accuracy was highest for cellulose, with median bias of −4% in test 1 and −6% in test 2, and worst for the high protein whey substrate in test 2, at −16% (Table 2, Fig. 2). The bootstrap analysis confirmed a bias for sunflower seeds ($p = 0.014$), bird feed ($p = 0.044$), and whey ($p < 0.001$), but not

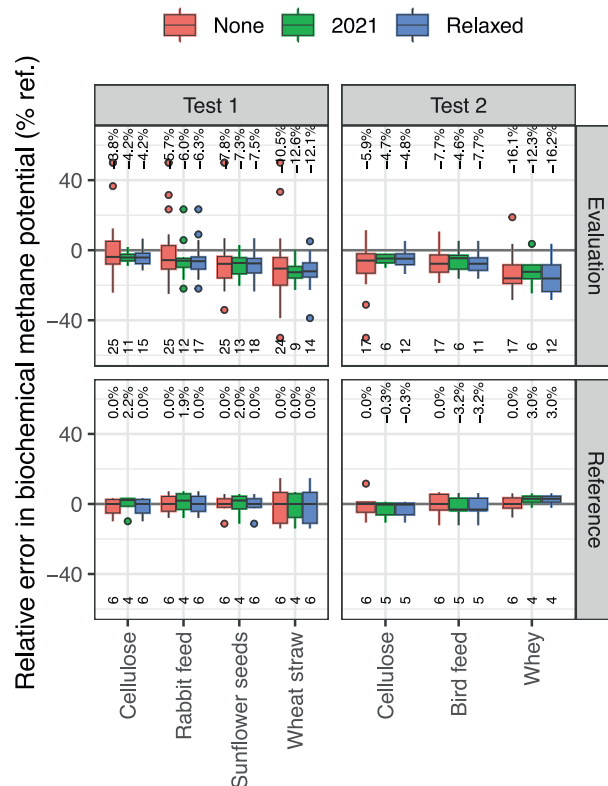


Fig. 2. Boxplot of submitted BMP accuracy by comparison to reference values (median BMP values from reference laboratories). The three colors show all observations (red, left), and results remaining after application of the 2021 criteria (Holliger et al., 2021a) (green, middle) or the relaxed criteria (blue, right). Extreme values above 50% were set to 50% for plotting only. Numeric values show the number of laboratories remaining in each group after application of validation criteria (bottom) and the overall error (median, top).

for cellulose ($0.1 < p < 0.22$) or wheat straw ($p = 0.13$), and the rabbit feed evidence was borderline ($p = 0.067$).

3.2. Total solids and volatile solids

In contrast to BMP, TS and VS measurements showed high reproducibility for both evaluation and reference groups. For TS, relative reproducibility standard deviation was better than 4% for all substrates. But there were a small number of TS and VS results with much higher errors; for TS the largest error was −17% of fresh mass, and −12% of TS for VS. Although consistent TS and VS measurements were not a challenge for most laboratories, these dry, ground substrates were not challenging and not representative of many typical biogas plant substrates. The possibility of higher error in TS and VS measurement during BMP testing remains.

3.3. Causes of error in BMP

With exceptions for some individual laboratories, the underlying causes of BMP measurement errors could not be fully explained. The only major errors that could be inferred conclusively from submitted data were in data processing and VS measurement by a single laboratory. A few laboratories reported explanations for improvements: a switch to gravimetric measurements (two laboratories), better mixing (one laboratory), and a switch to a better inoculum source (two laboratories).

Some of the least accurate BMP measurements, from laboratory E24 (Fig. 1), appeared to have errors related to data processing and VS measurement. More than half of the error in rabbit feed BMP submitted by E24 for test 1 (680 mL g^{-1}) was eliminated when moving from submitted BMP to BMP calculated with VS measured by the E24 laboratory. Most of the remainder was caused by the unusual VS value determined by this laboratory; BMP calculated using reference VS was within 10% of the reference value (Fig. 1). Similar (if less extreme) results were seen for other test 1 substrates for that particular laboratory. In test 2, the improvement in accuracy for E24 appeared to be related to somewhat more accurate calculations, although error remained high (Fig. 1).

Calculation errors were present in most evaluation laboratory results and, to a lesser degree, reference results. These errors contributed to observed variability in BMP but only accounted for a portion. Reproducibility standard deviation in test 1 for calculated BMP, reference VS, was substantially better than submitted BMP for only rabbit feed (20% vs. 26%) and wheat straw (25% vs. 33%). In test 2, submitted BMP had better reproducibility than calculated for bird feed (18% vs. 9%) and whey (17% vs. 13%), and in other cases differences were small. In parallel with better submitted BMP values, correspondence between submitted and calculated results improved in Test 2 (Fig. 1). Some of the original calculation problems were related to confusion in the submission process, e.g., submission of volume standardized to 0°C and 1.0 atm instead of as-measured volume, or VS given as a percentage of fresh mass and not TS. But these identified errors were corrected, and revised data were submitted; the remaining differences shown here remain unidentified. Median calculation error (including sign) was negligible: below 1% of reference values for all substrates for evaluation laboratories and below 0.05% for reference laboratories, suggesting there is no consistent positive or negative calculation error that laboratories typically make.

Laboratory E02 improved both the average BMP measurement accuracy and consistency among substrates from test 1 to 2 after switching from a manual volumetric method to a manual gravimetric method, based on the advice of project organizers. Laboratory E04 had a similar experience, switching from a manual manometric method. Gravimetric measurement may be more accurate than manual volumetric because it is unaffected by biogas leakage (Hafner et al., 2020). But these particular improvements do not indicate a problem with manual volumetric

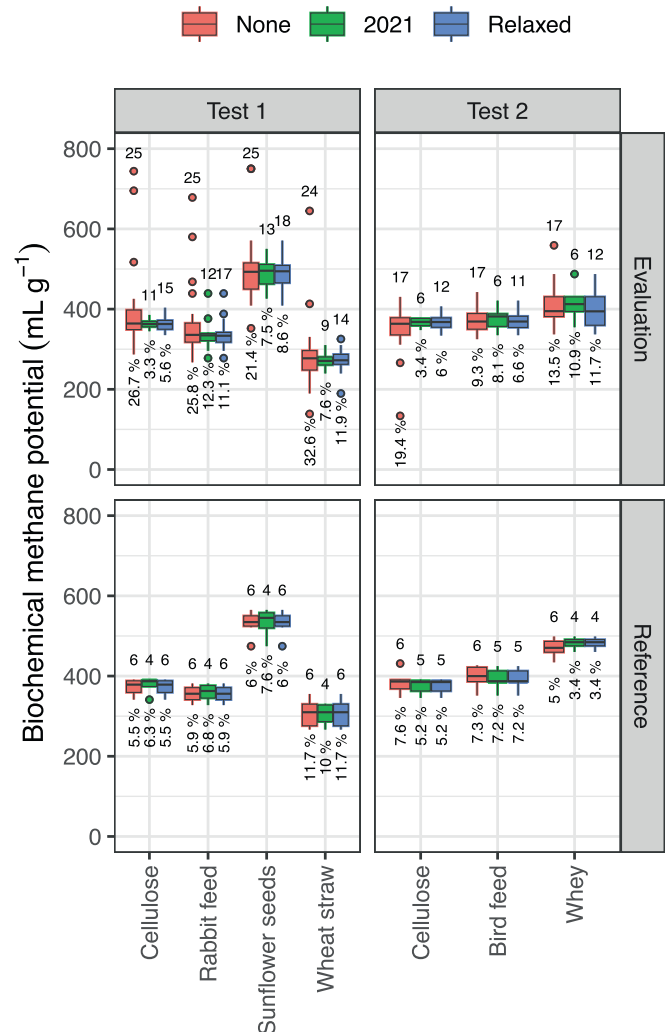


Fig. 3. Boxplot of submitted BMP values by substrate, test, laboratory type (evaluation and reference), and with two different sets of evaluation criteria applied. 2021 criteria include cellulose BMP limits of 340 and 395 mL g^{-1} and a relative standard deviation limit of 6%. The relaxed criteria have cellulose BMP limits of 330 and 415 mL g^{-1} and a relative standard deviation limit of 15%. Both include the 1% 3 d net duration criterion. Numeric values above the boxes show the number of laboratories remaining in each group, and values below show reproducibility standard deviation after exclusion.

measurement itself; other laboratories used it successfully.

Error in gas quantity measurement probably made a significant contribution to inaccurate results and low reproducibility. Earlier work has shown that this source can be quite significant with manual volumetric or manometric methods due to leakage or other, less clear mechanisms. Evaluation of manual manometric and volumetric results with gravimetric measurements in a large interlaboratory study showed that 20% error is not uncommon (Hafner et al., 2020). Automated commercial systems are not completely immune to these types of measurement errors. The extremely high test 1 BMP values from laboratory E01 were due to very high gas volumes which were measured with a commercial system, and the underlying cause of this unusual problem was not identified. But underestimation of biogas production is probably the more common case and may be a primary cause of the persistent negative bias observed in this work.

There was no clear evidence of consistent biases between the measurement methods. And reference laboratories provided evidence that very different methods can provide comparable results. Their methods included AMPTS®, other automated systems, manual volumetric and

manometric methods, and the gas density method, yet reproducibility standard deviation was low (Table 2). But lack of consistent bias in measurement method categories does not mean that biases do not exist in specific implementations of methods. This second source of error is almost certainly substantial, and the underlying causes of these errors likely vary. For example, laboratory E12 attributed their improved performance in test 2 to better mixing; in test 1 some substrate material remained attached to the side of the bottle just above the slurry, and was therefore most likely not digested.

Bias in BMP differed between substrates, suggesting that inoculum activity could be another primary cause of measurement error. The smallest bias for cellulose was consistent with high inoculum activity toward carbohydrates, while poorer accuracy in whey BMP implies that many inocula struggle to degrade protein completely or rapidly (Table 2, Fig. 2). The better performance by reference laboratories shows that suitable capable inocula exist. And results from two evaluation laboratories provide further evidence. Laboratory staff at both E14 and E22 attributed an improvement in BMP accuracy from test 1 to test 2 to a change in inoculum source. Laboratory E14 used a laboratory source for test 1 and supplemented it with full-scale digestate for test 2. Although a positive control substrate can highlight problems with inoculum performance, the use of cellulose will only help for carbohydrates. In fact, both of these evaluation laboratories submitted accurate cellulose BMP results for both tests while results for most of the complex substrates were 20–50% below reference results in test 1 (Figs. 1 and 2). It is also interesting that differences in reproducibility and accuracy among the substrates from the results as a whole are consistent with a general inoculum cause of BMP measurement error; differences between evaluation and reference laboratories in reproducibility and BMP values (accuracy) were smallest or even undetectable for cellulose, but clear and larger for the more complex substrates that were apparently difficult to degrade, particularly the lipid-rich sunflower seeds and protein-rich whey powder. Previous research has shown some effects of inoculum selection on BMP, but generally only small or even nondetectable effects (Koch et al., 2017a; Hülsemann et al., 2020), with some exceptions (De Vrieze et al., 2015). Biogas composition might give some indication of inoculum activity, with low CH₄ concentration indicating a stressed inoculum. Most values measured by those laboratories that did not remove CO₂ were typical, but a few were low, < 0.3 (volume fraction) for bottles with substrate and even < 0.1 for two blanks. However, there was no clear relationship between these results and BMP error.

3.4. Validation criteria

The 2021 validation criteria (Holliger et al., 2021a), with cellulose BMP limits of 340 and 395 mL g⁻¹, eliminated up to 15 of 25 evaluation laboratories in test 1, but also 2 of 6 reference laboratories (Fig. 3). For test 2, 11 of 17 evaluation laboratories were eliminated by 2021 criteria, and only 1 reference laboratory (2 for whey, because of a borderline duration). The main cause of elimination by application of the 2021 criteria was exceeding cellulose BMP limits (11 of 31 in test 1) but a surprising number of submissions did not meet the 1% net 3 d duration criterion (up to 6 of 31 for cellulose and wheat straw in test 1).

BMP reproducibility drastically improved for evaluation laboratories by application of the 2021 criteria, at least for substrates where the initial reproducibility standard deviation was high. For cellulose in test 1, for example, reproducibility standard deviation dropped from 27% to 3% after application of the 2021 criteria. But for the reference group, elimination of two the laboratories by application of the 2021 criteria provided no improvement; reproducibility standard deviation actually increased in some cases (Fig. 3). Interestingly, elimination of results based on the 2021 criteria did not consistently improve BMP accuracy. Median accuracy was generally stable, with only whey and bird feed improving (Fig. 2).

These results show that the 2021 criteria can eliminate grossly inaccurate BMP results and perhaps less extreme but still inaccurate

values, but the criteria may be too strict. Application of the relaxed criteria, with 330 and 415 mL g⁻¹ limits, support this interpretation. These criteria eliminated only one of the reference laboratories, fewer evaluation laboratories (up to 10 of 25 in test 1, 5 or 6 of 17 in test 2), and still provided similar accuracy (Fig. 2) and similar reproducibility (Fig. 3) as the 2021 criteria.

3.5. SMP curves

It has been proposed that specific methane production (SMP) curves can be used as a screening tool to identify problematic BMP tests, with a long “lag phase” with little or slow methane production or other trajectories differing from first-order responses indicating a problem (Koch et al., 2019). Curves of SMP calculated by project organizers from submitted raw measurement data in the present study show substantial diversity and major deviation from the first-order response in many cases, including lag phases above 20 d, with a median of 1.7 d (excluding cellulose, which typically has a lag phase and shows a median value of 3 d here). Almost half the laboratory × substrate combinations were flagged as problematic in the SMP curve evaluation, with nearly all of these coming from the evaluation laboratories (orange lines in Fig. 4). In general, reference laboratories showed curves that closely matched the ideal response (Koch et al., 2019), while many evaluation laboratory curves showed problematic results, even in test 2 (Fig. 4). Some evaluation laboratory SMP curves included declining values and even negative SMP, undulations, and other shapes inconsistent with a first-order response that imply problems with inoculum activity or BMP test execution (Koch et al., 2019).

The qualitative assessment of SMP curves showed clear value. For the two animal feeds, all extreme BMP values were associated with problematic SMP curves (Fig. 4). Results were similar for the other substrates, with few exceptions. The magnitude of relative error in submitted BMP value was clearly correlated with SMP curve status ($p < 0.002$ from a two-sample *t*-test for bird feed). Quantitative indicators of SMP curve shape showed somewhat less value. Error in submitted BMP was in fact correlated with NSE for a first-order model fit to the SMP curves but only for the evaluation laboratories ($p < 0.0001$ from a *t*-test on the model coefficient) with an increase in absolute error from around 10% to 30% as ME decreases from near 1.0 to 0.25.

Steindl et al. (2023) further refined the SMP screening concept and proposed that the SMP curve of blanks (with only inoculum) should have a lag phase < 1 d and relative standard deviation below 10% if BMP tests are to provide more reliable kinetic results. Here, however, BMP measurement error was generally not correlated with lag phase duration of substrates or blanks ($p > 0.07$). However, it does seem significant that not a single reference laboratory blank SMP curve showed long lag phases (0 with > 2 d lag phase and only 1 of 25 with > 1 d). Neither was there evidence of value in the initial methane productivity rate (normalized by inoculum VS mass) ($p > 0.5$ for evaluation laboratories and $p > 0.06$ for reference). Error in submitted BMP was, however, negatively correlated with NSE for the blanks ($p = 0.016$). But there were both observations with poor NSE and low BMP error, and a clear first-order response with high NSE but unusual BMP values with high error (Fig. 4).

3.6. Recommendations

The results from this interlaboratory study support some recommendations for laboratories, customers, and the research community. This work clearly demonstrates that experience matters, and that being receptive to advice from experienced staff can improve measurement quality, even without a laboratory visit. Laboratory staff should expect that measurement quality will improve with experience and seek out advice from more experienced laboratories. Given the diversity of methods and opinions, it is prudent to seek advice from multiple laboratories. Seeking advice and developing experience does not replace or

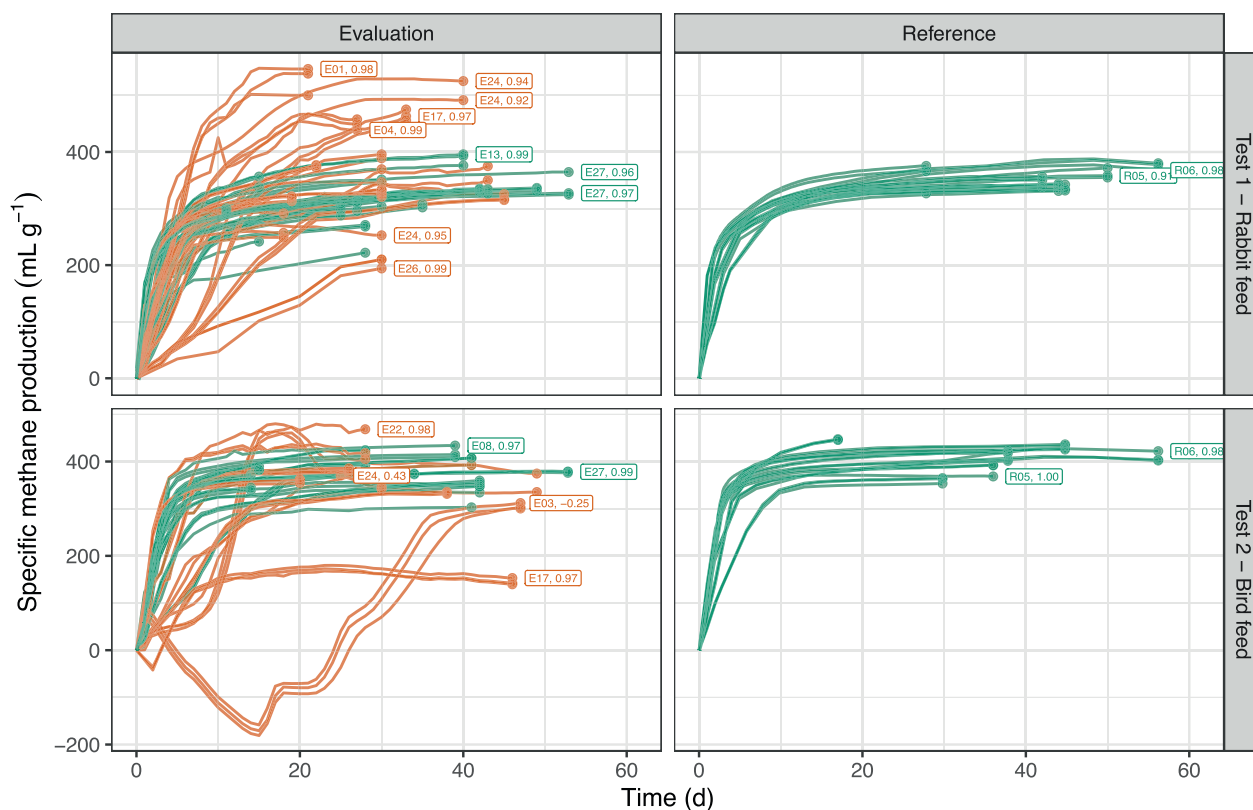


Fig. 4. Specific methane production (SMP) curves measured by participating laboratories for the animal feed substrates. Colors show visual qualitative evaluation results, with teal curves acceptable and orange flagged for obvious problems, including negative values, long lag phases, or shapes that otherwise differ from a first-order response. Selected replicate curves are identified, with labels showing laboratory key and Nash-Sutcliffe model efficiency (for first-order model with no exclusion of lag phase).

eliminate the need to follow available BMP measurement guidance, such as VDI (2016), VDLUFA (2011), or Holliger et al. (2021). However, a comparison between the present and earlier work suggests that simply following a common protocol is not sufficient. Neither Ribeiro et al. (2020) nor Hafner et al. (2020) found any improvement in BMP reproducibility after implementing a standardized protocol before a second round of BMP tests, while improvement was obvious in the present work.

Although preferences vary and all popular measurement methods for BMP can provide accurate results, laboratories using manual methods should consider including a gravimetric check of their results. The mass loss of bottles from the start to the end of a BMP test can confirm biogas production measured using manometric or volumetric methods, or reveal the presence of measurement errors (Hafner et al., 2020). A manual gravimetric method can circumvent the challenge of avoiding biogas leakage. Regardless, any BMP method should be based on methane production, and therefore biogas analysis or CO₂ removal is essential.

An effective inoculum is perhaps the most important component of a BMP test. Inocula should come from full-scale anaerobic reactors if possible (VDLUFA, 2011; VDI, 2016; Holliger et al., 2016). Although differences between different sources of effective inoculum may be minor, selecting a poor inoculum can undoubtedly result in underestimation of BMP (Koch et al., 2017b; Hülsemann et al., 2020).

BMP testing should always include evaluation of a positive control substrate and application of validation criteria. Microcrystalline cellulose has been used extensively (VDLUFA, 2011; VDI, 2016) and despite its limitations (Koch et al., 2020), can be effective in identifying problematic BMP tests when combined with validation criteria. Other substrates containing more than just carbohydrate will undoubtedly

provide a better test of the ability of an inoculum to degrade diverse substrates. Given the persistent negative bias shown in this work, and especially its variation among substrates, with the highest bias for whey, the development of more complex positive control substrates would be valuable. In the meantime, laboratories can consider checking for high conversion efficiency of food products (Koch et al., 2020) and participation in interlaboratory comparisons with complex substrates.

A perusal of recent published work on BMP shows that validation criteria are often omitted (e.g., García et al., 2023; Sillero et al., 2023; Dhulia et al., 2025) and that multiple sets are in use, including Holliger et al. (2016) (Patil and Sharara, 2022) and VDI (2006) (Gueri et al., 2025), both of which have been superseded, by Holliger et al. (2021a, b) and VDI (2016). Several papers have used these latest criteria as well. Results from the present contribution suggest that the 2021 criteria may be too strict. The biogas community might consider whether new criteria are needed. In any case, when validation criteria are applied, they should be clearly described.

Plots of SMP curves clearly have value (VDI, 2016; Koch et al., 2019; Hafner et al., 2020), and their inclusion in BMP test reports provided to customers, as well as in Supplementary Material of research papers on BMP, should become standard practice. This recommendation has been made previously, but as with validation criteria, uptake has been slow. The present contribution provides further evidence of the utility of SMP curves, and confirms that it is good practice to aim for curves that reflect reference laboratory results and to repeat tests that show substantially different responses (Fig. 4).

Data processing is not the main source of error in BMP measurement on average, but it can be the main source in some cases. As recommended earlier, this “low-hanging fruit” should be addressed by comparing calculated results to those obtained from a standardized

approach (Hafner et al., 2020). Free biogas software exists as an R package and web application (Hafner et al., 2018) and experienced laboratories may be willing to share spreadsheet templates or scripts for checking.

TS and VS were measured accurately by most laboratories, with results showing high reproducibility. However, as with calculations, there were cases where TS or VS measurement was a major source of error in BMP. And, as mentioned above (section 3.2), typical laboratory samples are more challenging than the dry homogeneous substrates used in the present study. Laboratories should test their TS and VS methods by processing substances with known values (e.g., microcrystalline cellulose should have both TS and VS close to 100%) or, better, through interlaboratory comparisons (Paterson et al., 2020).

Interlaboratory comparisons or more structured proficiency tests (Paterson et al., 2020) are useful for BMP testing in general, and both experienced and less experienced laboratories should participate in such projects regularly. These projects do not need to be large or complex; a simple substrate exchange with one or two laboratories within the same country can be useful. The substrates used in this present project and in earlier work (Raposo et al., 2011; Koch et al., 2020) can provide inspiration for complex yet stable substrates that are well-suited for these tests. The inclusion of some feedback or a troubleshooting discussion (Paterson et al., 2020), ideally between two tests, appears to be a promising way to identify and address problems, along with confirmation of improvement. The use of identical substrates in each test would facilitate this type of analysis but then introduces the possibility of bias due to participants knowing what to expect.

In summary, recommendations for biogas laboratories are: (1) work to accumulate BMP testing experience, (2) seek out method advice from more experienced staff, (3) use gravimetric methods to eliminate error from bottle leakage, (4) find or develop an effective inoculum, (5) regularly include a positive control substrate in BMP tests, (6) apply and report validation criteria, (7) compare BMP results to expected values, (8) check SMP curves of substrates and blanks and discard strange curves, (9) confirm the accuracy of data processing steps, and (10) participate in inter-laboratory comparisons or proficiency tests for BMP as well as TS and VS measurement.

Customers are recommended to ask biogas laboratories: (1) if they have recently participated in an interlaboratory test, (2) whether they regularly include a positive control, (3) whether they apply any validation criteria to each test, and (4) if the SMP curves have been evaluated for the test that included the customer's samples. Customers should not accept a simple "yes" or "no" but should ask laboratories to share results with them, including any interlaboratory report, positive control BMP, applied validation criteria and results, and SMP curves for the customer's substrates and positive controls in the same test.

4. Conclusions

Experience matters when it comes to BMP measurement, and participation in an interlaboratory comparison and evaluation such as the one described here can provide valuable experience and feedback to help improve BMP reproducibility. However, improving reproducibility may be easier than eliminating a small negative bias, which may require greater effort to identify sources of error. There is no simple way to guarantee that specific BMP results are accurate, but following existing guidelines and the recommendations provided in this paper would substantially reduce the magnitude of errors overall and in specific cases. Application of validation criteria based on a positive control substrate and verification of specific methane production curves are particularly effective in improving accuracy and reproducibility.

CRedit authorship contribution statement

Sasha D. Hafner: Writing – review & editing, Writing – original draft, Visualization, Validation, Supervision, Software, Resources,

Project administration, Methodology, Investigation, Funding acquisition, Formal analysis, Data curation, Conceptualization. **Estelle Leca:** Writing – review & editing, Visualization, Supervision, Resources, Project administration, Investigation, Conceptualization. **Cecilia Sambusiti:** Writing – review & editing, Supervision, Project administration, Funding acquisition, Conceptualization. **Sergi Astals:** Writing – review & editing, Resources, Investigation. **Olivier Azam:** Writing – review & editing, Resources, Investigation, Conceptualization. **Konrad Koch:** Writing – review & editing, Resources, Investigation. **Jing Liu:** Resources. **Henrik B. Møller:** Writing – review & editing, Resources. **Mihaela Nistor:** Writing – review & editing, Resources, Investigation. **Mario Olaya-Rincon:** Writing – review & editing, Visualization, Resources, Investigation. **Christine Peyrelasse:** Writing – review & editing, Resources, Investigation. **Alastair J. Ward:** Writing – review & editing, Resources, Investigation. **Bastien Zennaro:** Writing – review & editing, Methodology, Investigation. **Florian Monlau:** Writing – review & editing, Supervision, Project administration, Funding acquisition, Conceptualization.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgments

Funding for this project was provided by TotalEnergies Onetech R&D (Pau, France). This work would not have been possible without the time and effort of the staff at participating laboratories. The manuscript benefited substantially from detailed critical comments provided by two anonymous reviewers through multiple rounds of review.

Appendix A. Supplementary data

Supplementary data to this article can be found online at <https://doi.org/10.1016/j.biortech.2026.134136>.

Data availability

Anonymous measurement data from participating laboratories are publicly available at <https://github.com/AU-BCE-EE/Hafner-2026-IL-BMP>.

References

- De Vrieze, J., Raport, L., Willems, B., Verbrugge, S., Volcke, E., Meers, E., Angenent, L.T., Boon, N., 2015. Inoculum selection influences the biochemical methane potential of agro-industrial substrates. *J. Microbiol. Biotechnol.* 8, 776–786. <https://doi.org/10.1111/1751-7915.12268>.
- Dhulia, A., Ganguly, R., Sharma, A., Thakur, D., Boufadel, M.C., 2025. Investigation of BMP and energy generation potential from municipal solid waste under anaerobic condition. *Clean Techn. Environ. Policy* 27, 6793–6807. <https://doi.org/10.1007/s10098-025-03238-1>.
- Efron, B., Tibshirani, R.J., 1994. *An Introduction to the Bootstrap*. CRC Press.
- Elbeshbishy, E., Nakhla, G., Hafez, H., 2012. Biochemical methane potential (BMP) of food waste and primary sludge: influence of inoculum pre-incubation and inoculum source. *Bioresour. Technol.* 110, 18–25. <https://doi.org/10.1016/j.biortech.2012.01.025>.
- García, O.D.V., Neto, A.M.P., Moretto, M.R.D., Zaiat, M., Martins, G., 2023. Food-to-microorganism ratio as a crucial parameter to maximize biochemical methane potential from sugarcane vinasse. *Braz. J. Chem. Eng.* 40, 655–671. <https://doi.org/10.1007/s43153-022-00270-2>.
- Garfi, M., Martí-Herrero, J., Garwood, A., Ferrer, I., 2016. Household anaerobic digesters for biogas production in Latin America: a review. *Renew. Sustain. Energy Rev.* 60, 599–614. <https://doi.org/10.1016/j.rser.2016.01.071>.
- Gueri, M.V.D., Buratto, W.G., Nied, A., Schirmer, W.N., Frare, L.M., Furtado, A.C., 2025. Alkaline pretreatment of food waste and grass clippings for the enhancement of methane production in anaerobic digestion processes. *J. Mater. Cycles Waste Manage.* 27, 2652–2665. <https://doi.org/10.1007/s10163-025-02270-z>.

- Hafner, S.D., Koch, K., Carrere, H., Astals, S., Weinrich, S., Rennuit, C., 2018. Software for biogas research: tools for measurement and prediction of methane production. *SoftwareX* 7, 205–210. <https://doi.org/10.1016/j.softx.2018.06.005>.
- Hafner, S.D., Astals, S., 2019. Systematic error in manometric measurement of biochemical methane potential: sources and solutions. *Waste Manag.* 91, 147–155. <https://doi.org/10.1016/j.wasman.2019.05.001>.
- Hafner, S.D., Astals, S., Holliger, C., Koch, K., Nielsen, L., Refsahl, L., Weinrich, S., 2022. Assessing the value of kinetic results from biochemical methane potential tests: reproducibility from a large inter-laboratory study. *Clean. Chem. Eng.* 4, 100065. <https://doi.org/10.1016/j.clece.2022.100065>.
- Hafner, S.D., Fruteau de Lacroix, H., Koch, K., Holliger, C., 2020. Improving inter-laboratory reproducibility in measurement of biochemical methane potential (BMP). *Water* 12, 1752. <https://doi.org/10.3390/w12061752>.
- Holliger, C., Alves, M., Andrade, D., Angelidaki, I., Astals, S., Baier, U., Bougrier, C., Buffière, P., Carballa, M., de Wilde, V., Ebertseder, F., Fernández, B., Ficarra, E., Fotidis, I., Frigon, J.-C., Fruteau de Lacroix, H., Ghasimi, S.M.D., Hack, G., Hartel, M., Heerenklage, J., Sarvari Horvath, I., Jenicek, P., Koch, K., Krautwald, J., Lizasoain, J., Liu, J., Mosberger, L., Nistor, M., Oechsner, H., Oliveira, J.V., Paterson, M., Pauss, A., Pommier, S., Porqueddu, I., Raposo, F., Ribeiro, T., Rüsch Pfund, F., Strömberg, S., Torrijos, M., van Eekert, M., van Lier, J., Wedwitschka, H., Wierinck, I., 2016. Towards a standardization of biomethane potential tests. *Water Sci. Technol.* 74, 2515–2522. <https://doi.org/10.2166/wst.2016.336>.
- Holliger, C., Astals, S., Fruteau de Lacroix, H., Hafner, S.D., Koch, K., Weinrich, S., 2021a. Towards a standardization of biomethane potential tests: a commentary. *Water Sci. Technol.* 83, 247–250. <https://doi.org/10.2166/wst.2020.569>.
- Holliger, C., Fruteau de Lacroix, H., Hafner, S.D., Koch, K., Weinrich, S., Astals, S., et al. 2021b. Requirements for measurement and validation of biochemical methane potential (BMP). Standard BMP Methods document 100, version 1.9. Available online: <https://www.dbfz.de/en/BMP>. Accessed 1 February 2025.
- Hülsemann, B., Zhou, L., Merkle, W., Hassa, J., Müller, J., Oechsner, H., 2020. Biomethane potential test: influence of inoculum and the digestion system. *Appl. Sci.* 10, 2589. <https://doi.org/10.3390/app10072589>.
- Justesen, C.G., Astals, S., Mortensen, J.R., Thorsen, R., Koch, K., Weinrich, S., Triolo, J. M., Hafner, S.D., 2019. Development and validation of a low-cost gas density method for measuring biochemical methane potential (BMP). *Water* 11, 2431. <https://doi.org/10.3390/w11122431>.
- Kemausuor, F., Adaramola, M.S., Morken, J., 2018. A review of commercial biogas systems and lessons for Africa. *Energies* 11, 2984. <https://doi.org/10.3390/en11112984>.
- Koch, K., Lippert, T., Drewes, J.E., 2017. The role of inoculum's origin on the methane yield of different substrates in biochemical methane potential (BMP) tests. *Bioresour. Technol.* 243, 457–463. <https://doi.org/10.1016/j.biortech.2017.06.142>.
- Koch, K., Hafner, S.D., Weinrich, S., Astals, S., 2019. Identification of critical problems in biochemical methane potential (BMP) tests from methane production curves. *Front. Environ. Sci.* 7, 178. <https://doi.org/10.3389/fenvs.2019.00178>.
- Koch, K., Hafner, S.D., Astals, S., Weinrich, S., 2020. Evaluation of common supermarket products as positive controls in biochemical methane potential (BMP) tests. *Water* 12, 1223. <https://doi.org/10.3390/w12051223>.
- Nash, J.E., Sutcliffe, J.V., 1970. River flow forecasting through conceptual models part I—a discussion of principles. *J. Hydrol.* 10, 282–290.
- Oliveira, J.V., Costa, J.C., Cavaleiro, A.J., Pereira, M.A., Alves, M.M., 2022. Effect of endogenous methane production: a step forward in the validation of biochemical methane potential (BMP) tests. *Energies* 15, 4696. <https://doi.org/10.3390/en15134696>.
- Patil, P.S., Sharara, M.A., 2022. Impacts of sonication on biomethane potential (BMP) and degradation kinetics of pig lagoon sludge. *Biosyst. Eng.* 223, 129–137. <https://doi.org/10.1016/j.biosystemseng.2022.08.008>.
- Paterson, M., Oechsner, H., Tillmann, P., 2020. Inter-Laboratory Test: KTBL/VDLUFA-Proficiency Test Biogas. Chapter 8. In: Liebetrau, J., Pfeiffer, D. (Eds.), *Collection of Methods for Biogas*, second ed. DBFZ Deutsches Biomasseforschungszentrum gemeinnützige GmbH, Leipzig DE.
- Raposo, F., Fernandez-Cegri, V., De la Rubia, M.A., Borja, R., Beline, F., Cavinato, C., Demirer, G., Fernandez, B., Fernandez-Polanco, M., Frigon, J.C., Ganesh, R., Kaparaju, P., Koubova, J., Mendez, R., Menin, G., Peene, A., Scherer, P., Torrijos, M., Uellendahl, H., Wierinck, I., de Wilde, V., 2011. Biochemical methane potential (BMP) of solid organic substrates: evaluation of anaerobic biodegradability using data from an international interlaboratory study. *J. Chem. Technol. Biotechnol.* 86, 1088–1098. <https://doi.org/10.1002/jctb.2622>.
- R Core Team, 2025. R: A Language and Environment for Statistical Computing. R Foundation for Statistical Computing, Vienna, Austria.
- Ribeiro, T., Cresson, R., Pommier, S., Preys, S., André, L., Béline, F., Bouchez, T., Bougrier, C., Buffière, P., Cacho, J., Camacho, P., Mazéas, L., Pauss, A., Pouech, P., Rouez, M., Torrijos, M., 2020. Measurement of biochemical methane potential of heterogeneous solid substrates: results of a two-phase French inter-laboratory study. *Water* 12, 2814. <https://doi.org/10.3390/w12102814>.
- Rittmann, B.E., McCarty, P.L., 2020. *Environmental Biotechnology: Principles and Applications*, second ed. McGraw Hill, New York, NY.
- Seminario-Córdova, R., 2023. Latin America towards sustainability through renewable energies: a systematic review. *Energies* 16, 7422. <https://doi.org/10.3390/en16217422>.
- Sillero, L., Solera, R., Perez, M., 2023. Effect of temperature on biohydrogen and biomethane production using a biochemical potential test with different mixtures of sewage sludge, vinasse and poultry manure. *J. Clean. Prod.* 382, 135237. <https://doi.org/10.1016/j.jclepro.2022.135237>.
- Sluiter Amie, Bonnie Hames, Ruiz, R., Scarlata, C., Sluiter, J., Templeton, D., Crocker, D., 2008. National Renewable Energy Laboratory technical report NREL/TP-510-42618. National Renewable Energy Laboratory, Golden, CO, USA. Revision August 2012. <https://docs.nrel.gov/docs/gen/fy13/42618.pdf>.
- Steindl, M., Dandikas, V., Lichti, F., Höcherl, S., Koch, K., 2023. A comprehensive study on the consequences of substituting energy crops by alternative substrates for biogas production in Germany. *Renew. Energy* 219, 119541. <https://doi.org/10.1016/j.renene.2023.119541>.
- VDI, 2016. *Fermentation of Organic Materials: Characterisation of the Substrate, Sampling, Collection of Material Data, Fermentation Tests*, No. VDI 4630. Verein Deutscher Ingenieure e.V., Düsseldorf, Germany.
- VDI, 2006. *Fermentation of Organic Materials*. Verein Deutscher Ingenieure e.V., Düsseldorf, Germany.
- Verband Deutscher Landwirtschaftlicher Untersuchungs- und Forschungsanstalten (VDLUFA) (Ed.), 2011. Method 4.1.1 measurement of biogas and methane yields in fermentation tests. In: *Handbuch der Landwirtschaftlichen Versuchs- und Untersuchungsmethodik (VDLUFA-Methodenbuch)*, Vol. VII, Die Untersuchung von Böden, 4.Auflage, VDLUFA-Verlag, Darmstadt.
- Weinrich, S., Schäfer, F., Liebetrau, J., Bochmann, G., Paterson, M., Oechsner, H., Tillmann, P., 2018. Value of batch tests for biogas potential analysis: method comparison and challenges of substrate and efficiency evaluation of biogas plants. *IEA Bioenergy*.
- Xue, S., Song, J., Wang, X., Shang, Z., Sheng, C., Li, C., Zhu, Y., Liu, J., 2020. A systematic comparison of biogas development and related policies between China and Europe and corresponding insights. *Renew. Sustain. Energy Rev.* 117, 109474. <https://doi.org/10.1016/j.rser.2019.109474>.