Meta-analysis of methane yields from anaerobic digestion of dairy cattle manure.

Nicole D. Miranda[a,*], Ramon Granell[b], Hanna L. Tuomisto[c], Malcolm D. McCulloch[a].

[a] *Energy and Power Group, Department of Engineering Science, University of Oxford, Oxford, UK. E-mails: [nicole.miranda; malcolm.mcculloch]@eng.ox.ac.uk*

[b] *Oxford e-Research Centre, University of Oxford, Oxford, OX1 3QG, UK.*

[c] *European Commission, Joint Research Centre (JCR), Institute for Environment and Sustainability, Via Enrico Fermi 2749, 21027 Ispra, Italy.*

Abstract

This work presents a novel meta-analysis for methane yields from anaerobically digesting dairy cattle manure. A database is compiled from 115 articles (2,181 cases) by systematically assessing the academic literature. Y_{CH4} and secondary variables collected in the database describe performance, operation conditions of digester and manure composition. Inputs to the database are normalised to enable inter-study uni- and multi-variate analysis, according to different types of digesters.

Most cases (N=1,299) are reported for batch digesters and meta-analysis results in a mean Y_{CH4} of 0.230±0.016 m^3kg^{-1}. CSTR (N=241), results in mean of 0.190±0.016 m^3kg^{-1} (random-effect model). Other continuous digesters, such as fixed-film, plug-flow and UASB require further cases for the application of meta-analysis. For non-specified semi-continuous digesters (N=347), the mean calculated by the random-effect model is 0.204±0.032 m^3kg^{-1} (N=27).

With respect to multi-variate analysis of operation conditions, batch digesters form four clusters, while two clusters are identified for CSTR and three for semi-continuous digesters. For variables describing manure composition, batch digesters present eight clusters, while CSTR and semi-continuous digesters have insufficient cases. The strongest correlations found for these secondary variables (Spearman's rho: -0.685 and -0.696) result for pH and Y_{CH4} in batch and CSTR, respectively.

Maximum Y_{CH4} estimated by meta-analysis are proposed to replace the default IPCC values because they are based on a larger sample size and integrate updated literature. In particular, markedly higher Y_{CH4} are obtained for Asia, Middle East and Indian Subcontinent than those in the IPCC guidelines.

Keywords: Anaerobic digestion, biogas, methane yield, dairy, manure management.

*Corresponding author. Email: nicole.miranda@eng.ox.ac.uk

1. Introduction

This meta-analysis examines the methane yields (Y_{CH4}) found in the literature for anaerobic digestion (AD) of dairy cattle manure. The dairy sector is responsible for 4% of global anthropogenic greenhouse gas (GHG) emissions [1]. To reduce the carbon footprint of milk-producing farms, AD has been proposed as a mitigating technology. AD converts the volatile solids (VS) in cattle manure to biogas (CH_4 and CO_2 mixture) and digestate [2]. Consequently, GHG emissions are decreased by: i) avoiding high global warming CH_4 from manure management systems; ii) generating renewable energy that substitutes fossil fuels and; iii) producing digestate that replaces mineral fertilisers. In order to quantitatively predict these benefits, assumptions on the performance of AD processes are critical [3].

In this respect, Y_{CH4} is a metric widely utilised in mass and energy balances to calculate the CH_4 production. The Y_{CH4} corresponds to the CH_4 formed by unit of VS and is defined as follows:

$$Y_{CH4}[m^3 kg^{-1}] = \frac{CH_{4,output}}{VS_{input}} \tag{1}$$

Theoretical, modelled and experimental values of Y_{CH4} are available in the literature. Theoretical estimates commonly use the Buswell equation [4, 5] which is based on the chemical composition and stoichiometry of degradation reactions. However, actual Y_{CH4} are less than theoretical values because not all organic material converts to biogas (e.g. non-degradable solids, inhibition of reaction and/or generation of new bacteria). Thus, utilising theoretical Y_{CH4} can over-estimate CH_4 production for AD scenarios. For models of Y_{CH4} (e.g. ADM1 [6] and Linke et al. [7]), validations are based on experimental work with limited AD operation conditions. Therefore, increasing the sample size of experimental results can provide a more robust source of validation.

Although, empirical values of Y_{CH4} are available in multiple (and increasing) publications, it is challenging to integrate results differing in operation conditions, slurry compositions and units of measurements. This work addresses this by applying meta-analysis techniques to systematically review and numerically combine the research on Y_{CH4} in the dairy sector.

Meta-analysis has traditionally developed in the medical sciences but in recent years has been applied in the the bio-energy sector [8, 9]. For biogas generation from dairy farms, qualitative reviews are dominant [10–12]. Only two meta-analysis [13, 14] exist for AD of dairy manure, however, they are focused on different aspects of the process (i.e. microbial communities [13] and the effects over farm emissions [14]). Additionally, to the authors' knowledge, this is the first work to integrate such numerous cases from the academic literature in the field.

The aim of this meta-analysis is to quantitatively combine Y_{CH4} measurements to enable a comprehensive understanding of this metric with respect to operational and composition variables. For this purpose, two main steps are carried out: creating a specific database and statistically analysing the collected information. A database is compiled because no specific dataset is available for the entire literature reporting Y_{CH4} of AD in dairy farms. Research articles are found by using particular keywords and applying defined criteria (Section 2.1). The Y_{CH4} are extracted from the selected work and entered to the database together with secondary variables (i.e. specifying operation and/or manure composition). In order to enable numerical combination, units of measurements for all variables are standardised (Section 2.2). To analyse the variables compiled in the database, uni- and multi-variate tests are introduced in Sections 2.3-2.4. In the Results and Discussion section, key aspects of the database are first described (Section 3.1). Thereafter, Y_{CH4} are discussed in terms of digester type (i.e. batch, continuous or semi-continuous). Multi-variate structure between key variables related to the performance of different AD systems is examined. Correlations and clustering of principal components are discussed in terms of sources of similarities and variations. The discussion finally compares results of meta-analysis models applied over maximum Y_{CH4} to the default values of the National Guidelines for GHG emissions of the Intergovernmental Panel of Climate Change (IPCC [15], i.e. widely used to predict CH_4 emission from manure management systems).

2. Methods

This section first presents the methodology to develop the novel database of Y_{CH4} for AD of dairy cattle manure. Second, the mathematical background for the statistical analysis of the variables in the database is introduced. All statistical calculations are performed in R v.3.0.2.

2.1. Article selection

A systematic search of articles in the academic literature is carried out to compile the database of Y_{CH4} of dairy cattle manure and secondary variables. The selection of articles is based on the following four criteria:

i) scientific publications in the academic literature that have been peer-reviewed and are in the field of AD;

ii) primary sources;

iii) available work in journals in English to which the authors have access and;

iv) inclusion of experimental data of Y_{CH4} for AD of dairy cattle manure.

The first criterion is defined to ensure quality results in articles that potentially report Y_{CH4} values. The ISI Web of Knowledge platform [16] is used because it contains mainly peer-reviewed journals. The publication language is set as "English" and the publication year is not constrained. It should be noted that, although China has played an important role in the development of AD plants (e.g. by 2007 it had 26.5 million biogas plants installed [17]), articles written in Chinese are discarded by this criterion. Similarly occurs for publications from India [18]. This is significant constraint of the language criterion. However, as results will show, the selection of English articles in this work markedly contributes in bringing together a large sample size.

Keywords, logical connectors and combinations shown in Figure 1 are used for the search of articles. The keywords aim to identify technology presence, agricultural sector and the metric of interest for this meta-analysis.

—Figure 1 here—

From the search results, titles and abstracts are used for a first filtering step. They are assessed to determine if studies report mono-digestion of dairy cattle manure and are primary sources (i.e. the latter corresponds to the fulfilment of the second criteria). Secondary sources are excluded to avoid duplicated values in the database. Therefore, review articles are rejected in this step. In addition, articles from a same author are verified to not repeat cases in the database.

From the titles and abstracts, articles available to the authors are collected (i.e. third criteria). The final filter step consists of selecting papers that include quantitative and empirical results (in text, tables or figures) of Y_{CH4} for dairy cattle manure, thus, complying with the fourth criteria.

2.2. Database input

Y_{CH4} are extracted from the selected articles to compile the database. Values of of Y_{CH4} are adapted so that all cases correspond to cumulative CH_4 production per kilogram of VS. Y_{CH4} values are standardised to normal conditions (20 °C, 1 atm) and given in "$m^3 kg^{-1}$" (i.e. volume of methane produced per mass of VS fed to digester.

When articles report various Y_{CH4} (e.g. profiles across time or measurements at different temperatures), multiple cases are logged into the database. To record differences between cases, secondary variables are also entered to the database (when available).

Qualitative secondary variables are: author name, journal name, year of publication and location of study (country and/or continent). If the latter is not available, the location of the first author is used. For quantitative secondary variables, the first column of Table 1 shows those that are found most frequently in the literature for AD of dairy cattle manure, thus, used in this work. These variables specify AD operation conditions and chemical compositions of the undigested slurries (i.e. before entering the digester). The units of measurements for secondary quantitative variables are also presented in Table 1. If a variable is reported in different dimensions, underlying data from the original article is utilised to reconcile the units. If insufficient information is available when dimension conversion is required, the original units of measurements are recorded into the database, but the variable is excluded for numerical analysis.

—Table 1 here—

To ensure that comparable experiments are combined, the results are presented separately according to type of digester (batch, continuous and semi-continuous) and differences are discussed.

2.3. Uni-variate analysis

Having the variables standardised to the required unit of measurement, normality of the distributions is assessed by the Shapiro-Wilk test [19]. Inter-quartile ranges (IQRs) are reported to examine the dispersion of the variables of the database. Because of the expected non-normal distributions of Y_{CH4}, Wilcoxon Signed Rank test [20] is used to evaluate significant differences between distributions.

—Table 2 here—

In order to obtain novel effect sizes (i.e. Y_{CH4}) that synthesise the literature, meta-analysis models were applied to the distributions from the database. Fixed- and random-effect models are applied given the mean (M_i), standard deviation (s_i) and sample size (N_i) for each case "i". The calculations undertaken to evaluate the new Y_{CH4}

from this work, are summarised in Table 2. The main difference between the models is that the fixed-effect model considers that each case corresponds to a measurement of effect size from a population with unique mean (μ). The estimation of μ is reported as M. For this model, the variance (V_i) is calculated with respect to μ. The random-effect model, in contrast, assumes a distribution of the population, hence multiple μ_i. To account for this, the model measures the variability of M_i with respect μ_i and of μ_i with respect to μ.

2.4. Multi-variate analysis

To assess the overall effect of the secondary variables in Table 1 over the Y_{CH4}, correlations are examined by using the Spearman Rank test [21]. This test determines ($p<0.05$) if the relation between variables is monotonic (non-specific to linear relations). This test is chosen because it is a robust method in the presence of outliers.

Principal component analysis is combined with cluster-analysis to study the underlying structure of the operation conditions and slurry compositions. This approach is adopted from Brambilla et al. [22]. As a first step, the most reported variables describing AD systems in the literature are identified to undergo principal component analysis. This technique is a unsupervised pattern-recognition method that allows reduction of dimensionality for multivariate datasets [23]. The dimension reduction allows visualisation of the database and synthesis of the data collected from the numerous cases in the database. Results of first and second principal components are presented by score plots.

As a second step, different groups of cases in the database are analysed by applying the most used clustering algorithm (k-means [24]) over the first two principal components. K-means clustering depends on the initialisation, therefore, experiments are repeated over 100 initial random centroids for 2 to 10 clusters. The number of clusters is defined by selecting the best Davies-Bouldin Index (DB Index) score from the k-means output [24].

3. Results and discussion

This section presents the outcomes of the database development and statistical analyses in five subsections. The first subsection covers results from the application of filtering criteria, general overview of the typology of the papers that conform the database (e.g. year of publication and journal) and introduction to types of digester found in the literature. The second to fourth subsections report the Y_{CH4} for batch, continuous and semi-continuous AD-systems, respectively. Y_{CH4} and its relation to secondary variables are examined by uni- and multi-variate analysis for these different types of digesters. In the last subsection, the meta-analysis of maximum CH_4 potentials is discussed to improve those currently used by IPCC.

3.1. Database development

The search of papers using the word combination described in the methods section, outputs 888 publications. By assessing these papers against the selection criteria (i.e. peer-reviewed research articles in the field of interest), 115 articles (fraction equivalent to 13.0% of the initially found) are chosen to construct the database (publicly available on the website of the Energy and Power Group from the University of Oxford: epg.eng.ox.ac.uk/content/anaerobic-digestion). The causes for the high rejection, according to criteria, are shown in Figure 2a.

—Figure 2 here—

Within the papers rejected by the first criterion, 352 correspond to other feedstock; 43 co-digest dairy cattle manure with other organic waste, excluding specific Y_{CH4} values of interest; 38 are not research articles and are, for example, patents or book chapters and; 29 articles do not assess the production of CH_4 from AD, focusing on different technologies for manure-management systems (e.g. hydrogen production). All articles rejected by the second criteria correspond to review publications. The third criteria discards 14 articles that report titles and/or abstract in English, however, their main content was in a different language. In addition, 33 articles (from which 22 correspond to two journals), are not accessible to the authors. It should be noted that the database includes 45 scientific journals, thus, the third is not an important constraint. The fourth criteria rejects 178 papers that do not include numerical results on Y_{CH4} of AD process and 31 that use modelled or assumed (i.e. non-empirical) Y_{CH4} values.

Regarding secondary variables that describe the typology of the studies, the origin and year of publication of the 115 selected articles are shown in Figure 2b and Figure 3, respectively. It is observed that European and American work is dominant with 46 articles and 44 articles, respectively. For other continents markedly fewer articles are found to report empirical Y_{CH4} values. The categorisation of cases conforming the database according to world region is of particular interest when comparing results from this meta-analysis to the default maximum Y_{CH4} of the IPCC (see Section 3.5).

—Figure 3 here—

With respect to year of publication, the literature covering AD of dairy cattle manure is increasing (see Figure 3). Results indicate that from 2007 to 2014 an average of 10 publication per year reported Y_{CH4} of dairy cattle manure. Before 2007 only two articles per year in average contain experimentally-measured Y_{CH4} of this dairy waste. This increase is expected because, although AD is a mature technology (i.e. implemented for waste treatment since 1880s in Europe [2]), today it is being re-visited as a mean to reduce carbon emissions in the context of international agreements. In addition, it is observed that the increase in publication begins the same year of the global financial crisis (2007-8), which could be related to dairy farms seeking to diversify their production activities.

From the selected papers 2,181 cases are extracted and conform the database. To avoid duplicates, repeated cases found in five of the articles [25–29] are entered once. 2,161 of the cases in the database are selected from 95 articles. The remaining 20 articles report single cases from which 16 are identified as control experiments in co-digestion studies. Table 1 summarises secondary quantitative variables in the database according to digester type. It is observed that the cases in the database report different operation conditions and composition variables, according to the objectives of each article. For example, for batch and semi-continuous cases the time and hydraulic retention time, respectively, are specified in less cases than the temperature. However, for continuous digesters the opposite occurs. In order to improve the consistency in reporting the performance of AD systems in the research literature, it is recommended to follow standardised protocols such as the EPA Guidelines [30]. In consequence, the implementation of standardised protocols to report specifications on AD systems (and composition of inlet substrate) would allow the selection of a larger sample size for the multi-variate analysis of this work.

As presented in Table 1, 2,147 of the cases in the database can be classified as: batch, continuous or semi-continuous. The following sections analyse the Y_{CH4} and secondary quantitative variables separated in these categories. It should be noted that cases which report Y_{CH4} in units of measurements that cannot be reconciled to "m^3kg^{-1}", are omitted.

3.2. Batch digesters

Batch reactors are the most reported in the literature for the AD of dairy cattle manure with 1,299 cases. From these, 1,144 cases from 54 articles report Y_{CH4} in the units of interest (or can be converted to it). Digesters operated as batch reactors are loaded at the beginning of the experiment (or cycle) and only biogas is extracted during the AD process. The yields from these digesters result in non-normal distribution (Shapiro's P=

3.06x10^{-33}), median 0.124 m^3kg^{-1} and inter-quartile range (IQR) 0.235 m^3kg^{-1}. Low median is due to the presence of cases at low operation times, where the AD process is only beginning. In fact, 19 articles report profiles of Y_{CH4} against time, which indicate that multiple cases represent initial stages of CH_4 production.

To observe underlying structure of the Y_{CH4} from batch digester, results from principal component analysis and their subsequent cluster analysis are shown in Figures 4a-b for operational and composition variables, respectively. For the former, 42 studies report simultaneously Y_{CH4}, digester volume, temperature and operation time, yielding 1,104 cases. The first two principal components of Figure 4a represent jointly the variance fraction of 61.6%. The first principal component is highly dependent on Y_{CH4} and time of operation, while the contrary is observed for the second principal component.

—Figure 4 here—

The directions of the Y_{CH4} and time of residency vectors are similar. Thus, high Y_{CH4} are associated to longer digestion times, as expected. Higher correlation is also expected (Spearman's rho= 0.411, P-value= 2.2x10^{-16}), but not achieved due to the presence of overlapping profiles (i.e. see bi-plots of Y_{CH4} over time in Supporting Material).

Individual profiles of Y_{CH4} are observed to reach asymptotic maximums. This is due to decrease in the available organic material to be anaerobically digested. As time advances in batch digesters, bacterial population growth is limited by the availability of nutrients and therefore, these profiles are associated to Monod's equation [31].

With respect to the digester volume, the Y_{CH4} vector shows opposite direction (with respect to the first principal component axis), which can be attributed to variables presenting poor (and negative) correlation (Spearman's rho= -0.135, P-value= 4.68x10^{-6}). It is unexpected that the temperature vector is opposite to the Y_{CH4}, with respect to the first principal component axis. Much work (e.g. Umetsu et al. [32] and Kaparaju et al. [33]) have reported an increase of Y_{CH4} with raise of temperature, however, the batch digesters collected in the database show a weak and inverse correlation (Spearman's rho= -0.263, P-value= 2.2×10^{-16}). This could be explained by batch digesters achieving maximum Y_{CH4} independently of operation temperature as long as enough retention time is allowed.

Four clusters result from the k-means algorithm (DB Index= 0.25), as shown in Figure 4a. All clusters extend across the first principal component, while being less affected by the second principal component. The red cluster of Figure 4a, corresponds to small digesters (0.30-2.0 l) operating at the highest temperatures for batch digesters in the database (50 °C and 55 °C). This cluster is conformed by 73 cases from seven articles. The blue cluster of Figure 4a, represents large digesters (850 l) operating at 35 °C. This cluster of 34 cases is conformed by the data reported in only one article [34]. Both blue and red clusters correspond to the lower Y_{CH4} obtained for batch digesters with maximums 0.19 and 0.29 m^3kg^{-1}, respectively. The green and yellow clusters of Figure 4a correspond to the most common values for digester volumes (0.1 to 5.0 l) and temperatures (25 to 38 °C). These clusters are conformed by 653 cases (from 31 articles) and 344 cases (from 24 articles), respectively. The main difference between these groups is the time of reaction. In particular, the green cluster represents experiments carried out at lower retention times (up to 69 d) than those in the yellow cluster (up to 160 d). In agreement with this, the centroid of the green cluster corresponds to a lower Y_{CH4} ($\sim$0.05 m^3kg^{-1}) than that of the yellow cluster ($\sim$0.32 m^3kg^{-1}). An interesting case in the yellow cluster is that of Zeeman et al. [35] which, although outside the main group of cases, shows high Y_{CH4} (0.11 m^3kg^{-1}) for low operation temperature (25 °C).

For secondary variables describing initial compositions, only 4 studies report simultaneously Y_{CH4}, pH, DM content, VS content and COD values, yielding 108 cases. The first two principal components of these variables are shown in Figure 4b and represent the variance fraction of 74.8%. The variables that most affect the first principal component (variance fraction of 51.8%) are VS, DM and COD. These variables relate to the available organic matter to be digested, thus, their vectors in Figure 4b are in similar direction. DM and COD present weak negative correlations with respect to Y_{CH4} resulting in Spearman's rho of -0.361 (P-value= 6.3x10^{-18}) and -0.261 for COD (P-value= 2.5x10^{-9}), respectively. VS shows a moderate monotonic function (Spearman's rho= -0.457 and P-value= 1.6x10^{-27}). The second principal component is more affected by pH and Y_{CH4}. These vectors show a strong negatively correlation (Spearman's rho= -0.6960, P-value= 1.0x10^{-97}), which is reflected in the opposite direction of their vectors in Figure 4b. This is an expected result because high pH content (e.g. ammonia in the slurry) can inhibit the methanogenic activity [36]. Eight clusters result (DB Index= 0.13) from the k-mean applied to the principal components of composition variables for batch digesters. The cases align in diagonals which could suggest that other cluster algorithms are preferable (e.g. density-based). The particular large amount of clusters for this bi-plot could be due to overlapping data in each diagonal which can affect the DB Index optimisation [37]. The lines which arise from the cases are parallel with slope value -1.56. In the case of the yellow and light blue clusters, these align having pH 6.8, DM mass fraction 1.9%, VS dry mass fraction 62.7% and COD 21 $kg \cdot m^{-3}$. The Y_{CH4} for these clusters ranges from 0 to 0.33 m^3kg^{-1}. Their intercept is 1.17. The orange, blue and red clusters (bottom right of Figure 4b) are aligned with an intercept -3.09. The cases in these clusters, have fix value for pH (6.5), DM (mass fraction 2.3%), VS (dry mass fraction 74.7%) and COD (30.3 $kg \cdot m^{-3}$). The red cluster corresponds to a case from the work of Pandey et al. [38], who attribute the high results to differences in the inocula utilised (i.e. granular sludge from digester processing potato starch). However, it is considered to be an outlier in this meta-analysis because it reports 1.67 m^3kg^{-1} for Y_{CH4}, markedly exceeding the theoretical value (0.469 m^3kg^{-1} by Buswell's equation [39, 5]).

The light green cluster has pH 6.9, DM mass fraction 5.3%, VS dry mass fraction 73.6% and COD 70.3 $kg \cdot m^{-3}$, with Y_{CH4} ranging from 0.26 to 0.33 m^3kg^{-1}. The parallel that crosses this cluster has an intercept -5.78. The final cluster corresponds to a single case from Pabon-Pereira [40] (dark green cluster in Figure 4b). This has pH 6.9, DM mass fraction 10.5%, VS dry mass fraction 86.7% and COD 29.2 $kg \cdot m^{-3}$ (Y_{CH4} range= 0.14 to 0.23 m^3kg^{-1}). If the same slope was repeated for this case, the intercept is -10.6. Overall, an increase in the organic matter content of the slurry is observed according to the first principal component. However, the increase of these variables presents no correlation with the ranges of Y_{CH4} in each aligned group. This also is observed by the low correlations (i.e. Spearman rho values) between Y_{CH4} and DM, VS and COD.

3.3. Continuos digesters

Continuous and semi-continuous digesters are less found in the literature in comparison to batch systems. The Y_{CH4} distributions for these types of digesters are shown as green box-plots in Figure 5. The orange box-plot correspond to distribution of Y_{CH4} which are jointly reported with N_i and s_i, thus, can be used for meta-analysis. The results from meta-analyses models are also presented in Figure 5 (fixed-effect model: red circles and random-effect model: yellow circles).

—Figure 5 here—

The database includes 335 cases for continuous digesters (see Table 1), from which 331 (from 38 articles) are used because Y_{CH4} can be expressed as "m^3kg^{-1}". For these digesters, dairy cattle manure is fed continuously into the reactor while streams of biogas and digestate (products) are removed. Specific subcategories found for continuous digesters, according to reactor operation are (see Figure 5): continuously-stirred tank reactors (CSTRs), fixed-film reactors, plug-flow reactors and up-flow anaerobic sludge blanket.

There are 241 cases (from 35 articles) in the database that report Y_{CH4} from CSTR. The distribution of all Y_{CH4} for this digester type (green boxplot in Figure 5) is non-normal (Shapiro's P= 2.91x10^{-5}) and has a median 0.141 m^3kg^{-1}. As expected this distribution is significantly lower (Wilcox's P=2.97x10^{-16}) from the B_0 of batch digesters. This is due to significantly lower hydraulic retention times (HRTs) in continuous digeseters compared to residency time in batch digesters (Wilcox's P< 2.2x10^{-16}). From this distribution, 12 cases report Y_{CH4} with N_i and s_i (orange box-plots in Figure 5). This subgroup is non-significantly different from the distribution of all Y_{CH4} reported for CSTR (Wilcox's P=0.143). Therefore, meta-analysis models applied over this subgroup are representative of the literature data. The meta-analyses models applied over the latter cases results in 0.179 ± 0.001 m^3kg^{-1} for fixed-effects and 0.190 ± 0.016 m^3kg^{-1} for random-effects.

Multi-variate analysis is carried out for CSTRs because these are the most numerous cases reported for continuous digesters. Score-plots for operation variables associated with CSTR digesters are shown in Figure 4c. From 23 articles which report all operation variables of Table 1, 173 cases are extracted and used for principal component analysis. The first principal component of Figure 4c corresponds to the variance fraction of 43.2% and is highly dependent on Y_{CH4}, OLR and volume. Contrary to batch digesters, the volume and temperature vectors are in the same direction as the Y_{CH4} with respect to the first principal component. These both variables show a strong positive correlation with respect to Y_{CH4} (Spearman's rho= 0.71 for volume and; rho= 0.69 for temperature). The second principal component represents the variance fraction of 26.8% and is mostly affected by the HRT and OLR. A positive moderate correlation (Spearman's rho= 0.435, P-value: 2.7x10^{-16}) is found, between the Y_{CH4} and HRT. At lower HRT (< 40 d), Y_{CH4} exhibits a larger range (0.002 to 0.333 m^3kg^{-1}), while presenting more stable values thereafter. These observations are similar to the asymptotic profiles for Y_{CH4} and retention time in batch digesters (see biplots of Supporting Material). With respect to the OLR, the Y_{CH4} vector is in opposite direction having a Spearman's rho -0.45 (P-value= 5.2×10^{-11}). This is an expected result because increasing OLR (i.e. lowering the retention times) can wash-out bacterial cultures reducing the digester performance [41].

Two clusters are output from the application of k-means to the principal components of Figure 4c (DB Index= 0.37). The blue cluster is conformed by 95 cases from 16 articles. This cluster groups cases in a larger range of HRT (1 to 128 d) and OLR (0.12 to 34.4 $kg \cdot m^{-3} d^{-1}$), compared to the green cluster (HRT: 8 to 69 d and; OLR: 0.12 to 6.6 $kg \cdot m^{-3} d^{-1}$). The digesters operating within the blue cluster also have lower temperatures (10 to 55 °C) and result in lower Y_{CH4} (maximum 0.220 $m^3 kg^{-1}$). For the green cluster, 78 cases (from 11 articles) are observed in Figure 4c. The Y_{CH4} from cases in this cluster range from 0.115 to 0.333 $m^3 kg^{-1}$. In general, these cases operate at higher temperatures (35 to 55°C) and in larger digesters (up to 1,500 l) than those in the blue cluster (up to 230 l).

With regards to the composition variables (see Table 1), it is found that there are insufficient cases of CSTR digesters in the literature for dairy cattle manure to perform principal component analysis. However, there are enough cases for bi-variate analysis between Y_{CH4} and each composition variable (see Supplementary Material).

The variables most correlated with Y_{CH4} are found to be the pH (Spearman's rho= -0.685, P-value= 2.0×10^{-15}) and nitrogen content (i.e. TKN, Spearman's rho= 0.605, P-value= 7.8×10^{-13}) of untreated manure. The strong correlation between pH and Y_{CH4} is in agreement with the findings for batch digesters. However, high correlation between TKN and Y_{CH4} is unexpected because higher nitrogen in ammonia has been reported to inhibit AD [36]. For variables related to organic material in the input slurry, only weak correlations are found with respect to Y_{CH4}: DM (Spearman's rho= 0.188, P-value=0.018), VS (Spearman's rho= -0.338, P-value: 2.5x10^{-5}) and COD (Spearman's rho= 0.48, P-value= 1.9x10^{-8}).

Fixed-film reactors are the second most numerous continuous digesters collected in the database (N=81, from five articles). The Y_{CH4} for this digester type (green boxplot in Figure 5) is not normally distributed (Shapiro's P= 3.3x10^{-14}) and has median 0.103 m^3kg^{-1}. From these Y_{CH4}, eight cases from the work of Lo et al. [27] are the only to report N_i and s_i (orange box-plot Figure 5). The fixed-effect meta-analysis model applied over these eight cases results in Y_{CH4} mean of 0.017 ± 0.001 m^3kg^{-1}. For the random-effect model, the mean Y_{CH4} is 0.018 ± 0.006 m^3kg^{-1}. However, the means obtained by meta-analysis should be used with caution as the distribution of the eight cases in the subgroup is significantly different (Wilcox's P= 3.94x10^{-4}) from the Y_{CH4} in the 81 cases for all plug-flow digesters in the database.

For the remaining continuous digesters, insufficient cases are collected for the application of meta-analysis models. A single case is found to report experimental Y_{CH4} of dairy cattle manure in units of interest for plug-flow reactors (PFRs, N=1). Thus, the latter is ommited from Figure 5). For up-flow anaerobic sludge blanket (UASB) digesters, there are eight cases which report experimental Y_{CH4} and correspond to the work of Rico et al. [42]. The distribution of these digesters is normal (Shapiro's P= 0.58) and has median 0.484 m^3kg^{-1} (see Figure 5). The mean of these cases, exceeds the theoretical value (0.469 m^3kg^{-1} by Buswell's equation [39, 5]). However, they do not present a significant difference (Wilcox's P-value=1). Therefore, differences in Y_{CH4} could be attributed to chemical compositions or longer biomass retention time (i.e. bacterial consortia and organic substrate).

3.4. Semi-continuos digesters

For semi-continuous digesters the input and/or output of digesters are discontinuous. There are 513 of these cases identified in this work (Table 1) and 443 of them (from 26 studies) are used.

One type of semi-continuous digester corresponds to AD ponds. These usually accumulate manure periodically but are discharged sporadically (e.g. when the produced bio-fertiliser is needed). There are 96 cases identified in the database for lagoon digesters from three articles. The high variability in loading/discharging rates and operation temperature results in a large IQR of 0.214 m^3kg^{-1} (see Figure 5). The Y_{CH4} for lagoon digesters presents a non-normal distribution (Shapiro's P=2.84x10^{-6}) and median 0.345 m^3kg^{-1}. This median is not significantly different than the theoretical Y_{CH4} (Wilcox's P-value=0.422). The higher Y_{CH4} of the lagoon digesters distribution, has been attributed by other authors [43–45] to degradation of accumulated organic matter on the bottom of ponds. Thus, more VS are available to be converted to CH_4, than the measured at the inlet to the AD ponds. From the above, it is suggested that for future work, measurements of VS fed to lagoon digesters are standardised.

From the 96 cases, only two (from two articles [43, 44]) can be used for meta-analysis models (orange box-plot in Figure 5). The Y_{CH4} of these cases is non-significantly different from the distribution of all lagoon digesters (Wilcox's P=0.297). The fixed effect model results in a mean 0.400 ± 0.057 m^3kg^{-1} (see red circles of Figure 5) . This is in great agreement with the case from Safley et al. [43], because the sample size of Y_{CH4} from that case is 93, whereas, the case from Park et al. [44] has a sample size of seven. For the random-effect model a mean of 0.440 ± 0.160 m^3kg^{-1} is obtained.

There are 347 cases in the database (from 23 articles) that include Y_{CH4} from other non-specified semi-continuous digesters. These cases are grouped and also shown in Figure 5. Their distribution is non-normal (Shapiro's P<2.2x10^{-16}) with median 0.173 m^3kg^{-1}. A subgroup of 27 of these cases (from five articles) are utilised for meta-analysis models. For the fixed-effect model a mean 0.176 ± 0.001 m^3kg^{-1} is obtained and for the random-effect model 0.204 ± 0.032 m^3kg^{-1}.

There are 98 cases from 15 articles in the database that specify all operation variables of Table 1 and Y_{CH4}, for semi-continuous digesters (non-specified, i.e. anaerobic digesters labelled as ponds are excluded). These cases undergo principal component analysis and the resulting score-plots are presented in Figure 4d. With the excpetion of the volume vector, it is observed that according to the variables directions (red arrows of Figure 4d), the semi-continuous digesters are in closer agreement to the batch digesters than the continuous digesters. In fact, it is observed that the studies utilising semi-continuous digesters are classified as such because of the feeding procedure, however, the overview of many systems corresponds to dynamic states, rather than stationary (i.e. continuously operated digesters).

The first principal component represents the variance fraction of 43.7% and the second 27.8%. The first principal component is highly dependent on OLR and Y_{CH4} and exhibit a weak correlation (Spearman's rho=-0.262; P-value=1.0x10^{-5}). The second principal component is mostly affected by the volume and temperature. The direction of the Y_{CH4} vector is opposite (along the first principal component axis) to the digester temperature. This shows a weak and inverse correlation which is further indicated by the Spearman's rho value -0.162 (P-value= 2.6x10^{-3}). The volume of digesters also presents a weak negative correlation to the Y_{CH4} (Spearman's rho=-0.250; P-value=2.5x10^{-6}). Regarding the HRT and Y_{CH4} vectors, these have similar direction, with respect to the first principal component. These variables exhibit a moderate positive correlation (Spearman's rho= 0.242, P-value= 5.1x10^{-6}) and follow similar patterns than those observed for batch digester (see Supporting Material).

Three clusters are found (DBI=0.33) for semi-continuous digesters and presented in Figure 4d. The red cluster comprises 28 cases from three different articles of Lo et al. (see online database). It is observed that all cases in the cluster operate at thermophilic conditions (55 °C) and volume 3 l. The differences in principal components for this cluster are due to changes in OLR (2.07 to 35 $kg \cdot m^{-3} d^{-1}$) and HRT (1 to 15 d). The Y_{CH4} range is the lowest among those in Figure 4d, from 0.011 to 0.161 $m^3 kg^{-1}$. The blue cluster comprises 57 cases from 13 articles. These cases exhibit higher Y_{CH4} than the red cluster, with centroid at approximately 0.262 $m^3 kg^{-1}$. This is attributed to higher HRT (1 to 60 d) and, in turns, lower OLR (0.4 to 8.87 $kg \cdot m^{-3} d^{-1}$). The semi-continuous digesters of the blue cluster operate at the lowest range of temperatures (20 to 37°C), which is consistent with psychrophilic and mesophilic conditions. The volumes of the digesters in this cluster covers a large range (between 0.04 L and 300 L), suggesting that this operation variable is not affecting the principal components describing AD performance. The green cluster of Figure 4d corresponds to 13 cases from two articles. The Y_{CH4} are the highest found for the semi-continous digesters, with centroid at approximately 0.77 $m^3 kg^{-1}$. This exceeds batch performances, which could suggest the presence of lagoons which, as discussed previously (see Figure 5), have been identified to have higher performance due organic material accumulated in the bottom of digesters. The OLR (0.5 to 2.9 $kg \cdot m^{-3} d^{-1}$) and HRT (10-20 d) is not markedly different from the other clusters. The temperatures of the digesters in the green cluster are between 35 and 45 °C (i.e. mesophilic conditions). This suggests that temperature is a key factor in the separation of the three clusters. As for continuous digesters, there are insufficient cases of semi-continuos reactors in the database to perform principal component analysis over composition variables of interest. Considering bi-variate analysis some moderate correlations can be found between Y_{CH4} and secondary variables describing initial composition of the dairy cattle manure. The strongest correlations between Y_{CH4} and a secondary variable is with COD (Spearman's rho= -0.542, P-value= $2.2\text{x}10^{-16}$) and TKN (Spearman's rho= -0.485, P-value= $1.6\text{x}10^{-13}$). This interesting result supports the well-known work of Angelidaki et al. [36], which reports that an increase in ammonia (containing nitrogen) can suppress CH_4

production in AD of livestock waste. However, it should be noted that Sterling et al. [46] have reported that only large increases in ammonia can inhibit AD, while gradual increases could enhance the CH_4. The latter is not evidenced in the results of this work. The other composition variables of interest are less correlated or weakly correlated to Y_{CH4}: pH (Spearman's rho= -0.060, P-value= 0.41), DM (Spearman's rho= -0.354, P-value= 2.8×10^{-9}) and VS (Spearman's rho= -0.284, P-value= 3.5×10^{-6}).

3.5. Maximum methane potentials

This section considers the experimental Y_{CH4} (B_0) of dairy cattle manure (i.e. obtained from batch essays). These values are interesting to study separately from other cases, because they can be utilised to quantify energy output and environmental benefits of AD systems. Much work is based on the methodology described by Owen et al. [47] to experimentally determine B_0. However, it should be noted that there is little agreement in the use of this or any standard protocol [48]. This results in difficulties to identify the differences between AD systems. In particular, an essential issue to consider when comparing B_0 from different AD systems is to report the temperature and pressure at which the experiments are carried out (i.e. as reported by Marti-Herrero et al. [49]). The B_0 from the database are used to obtain novel effect sizes by meta-analysis models . Results are compared against the B_0 default values from the IPCC 2006 methodology [15], in accordance to the world regions defined in the latter.

The IPCC Guidelines provide two default values of B_0 from AD of dairy cattle manure. For Europe, Oceania and North America the default B_0 is 0.24 m^3kg^{-1} and is based on the work of Morris 1976 [50, 15]. For other regions of the world the default value 0.13 m^3kg^{-1} is based on the work of Safley et al. (1992, [51, 15]). A revision of these values is needed because they only consider a reduced amount of literature (from 1976 and 1992) and exclude the last thirteen years of research which, according to the findings of this work (see Figure 3), accounts for the fraction of 87.5% of published articles in the field. Furthermore, the B_0 of the original work is based on the diet of dairy cattle which are also in need of update.

The B_0 values from the IPCC are shown as blue circles in Figure 6. It should be noted that these default Y_{CH4} are not considered in the subgroup for meta-analysis because the original sources [51, 50] are published in books (i.e. not research articles as required by the first criterion).

—Figure 6 here—

From the database, 799 cases (from 42 articles) are identified as either biochemical potential tests or maximum Y_{CH4} determined in similar batch conditions (i.e. measuring cumulative CH_4 production until no increase is observed). For articles reporting Y_{CH4} against time, the maximum yield of each profile is extracted. In the 42 articles, 143 cases are identified as B_0. The distributions of B_0 are represented by the green box-plots of Figure 6. For all the regions defined in the IPCC Guidelines experimental results (in comparable units of measurements) are found for B_0 (see number of cases in Table 3), except Oceania and Latin America (i.e. omitted in Figure 6). As expected, the results for B_0 are significantly higher than the Y_{CH4} distribution for all batch cases (Wilcox's P= 9.67×10^{-10}).

—Table 3 here—

It is observed in Figure 6, that the B_0 values from the IPCC with corresponding error bars of 15%, are within the distribution of B_0 found in the database. In fact, only the region comprising Asia, Middle East and Indian Subcontinent shows a significant difference (Wilcox's P-values in Table 3) in B_0 distribution with respect to the IPCC value. This could be attributed to developments in feeding of dairy cattle in that region, because the IPCC default values are based on the dairy diets prior to 1992 [51]). Similar observations are found for African experiments (i.e. the fraction of cases which is greater than the default IPCC value is 60%). However, the sample size is not large enough (N=5) to prove significant difference. Europe and North America present the largest samples of B_0 (see Table 3), as expected from the origin of the articles identified in Section 3.1.

Although most regional B_0 values are non-significantly different from those in the database from the IPCC default values, results indicate that different means and error bars would better represent the findings of the literature. Therefore novel effect sizes and confidence intervals are examined by applying fixed- and random-effect models of meta-analysis.

The meta-analysis models require cases which report B_0 together with specific N_i and s_i. For this 53 cases from 21 studies are selected. The distributions of this subgroup of B_0 is shown in Figure 6 (orange box-plots). Results show that the means of these distribution differ with the IPCC default value between 0.01 and 0.13 m^3kg^{-1}. In particular, the subgroup distributions for the Asia, Middle East & Indian Subcontinent Region is again the only significantly different ($p<0.05$) from the IPCC value (i.e. as for the green box-plot of Figure 6, discussed previously).

The results of fixed-effect and random-effect models are also shown in Figure 6 in red and yellow circles, respectively. The means obtained by the random-effect models are found to be in best agreement with the distributions of the subgroup of B_0 cases (i.e. case reporting N_i and s_i), with differences from 0.001 to 0.026 m^3kg^{-1}. The means for this model are 0.230 ± 0.016, 0.230 ± 0.047, 0.230 ± 0.020, 0.260 ± 0.060 and 0.110 ± 0.022 m^3kg^{-1} for all regions, Asia/Middle East & Indian Subcontinent, West/East Europe, North America and Africa, respectively. For the fixed-effect model the means are 0.198 ± 0.001, 0.220 ± 0.003, 0.195 ± 0.001, 0.280 ± 0.011, 0.100 ± 0.009 m^3kg^{-1}, for the same regions mentioned above, respectively. This results in differences from 0.010 to 0.040 m^3kg^{-1} between the means of the fixed-effect model and those of the subgroup of B_0 cases. The random-effect model is expected to show better results because it assumes a distribution of B_0, rather than a fix value. This is more representative of the high heterogeneity of the samples, which arise from different experimental set-ups. Therefore, it is preferable to adopt the results of B_0 from the random-effect model to estimate the potential CH_4 production from AD of dairy manure.

Compared to the IPCC default values, the random-effect meta-analysis model provides more robust values of B_0 (i.e. based on larger sample size, more updated cases and considers the heterogeneity of cases). Therefore, the means and confidence intervals of B_0 obtained by this model are suggested to replace those in the current IPCC Guidelines. Furthermore, it is recommended to apply this meta-analysis model for the estimation of B_0 of other digestible feedstock. This would provide better IPCC-default B_0 values, with stronger evidence-base (i.e. larger sample integration).

4. Conclusions

This work presents a meta-analysis of methane yield of dairy cattle manure. A novel database is compiled from 115 selected articles (2,181 cases). The systematic review of the literature shows that the report of operation conditions and slurry composition does not generally follow standardised protocols. The latter are strongly recommended in order to provide insights of differences in performance of AD systems.

Meta-analysis models are applied over the database according to digester type (batch, continuous or semi-continuous). This methodology enables numerical combination and synthesis of experimental results to date.

Multi-variate analysis identifies clusters for Y_{CH4} and operation conditions (i.e. temperature, volume, time/HRT, OLR) or feedstock composition (i.e. pH, DM, VS, COD, TKN). This can contribute to the construction of AD scenarios by providing ranges of variables that are commonly grouped.

Meta-analysis applied over maximum Y_{CH4} is a reliable form to correct the out-of-date IPCC default values. This method has a greater sample size (i.e. based on more articles than the current IPCC Guidelines). Consequently, this work can contribute to more robust estimates of energy production and greenhouse gas mitigation of AD systems in the dairy sector.

Supporting Material

Supporting information containing biplots for Y_{CH4} against secondary quantitative variables and DB Index results is available online.

Acknowledgements

The authors would like to acknowledge Becas Chile (call of 2010) for funding the first author of this work.

References

[1] Food and Agriculture Organization of the United Nations, Animal Production and Health Division. Greenhouse Gas Emissions from the Dairy Sector a Life Cycle Assessment. Rome: FAO; 2010.

[2] Deublein D, Steinhauser A. Biogas from Waste and Renewable Resources. 2nd ed. Weinheim, Germany: Wiley; 2010.

[3] Berglund M, Barjesson P. Assessment of energy performance in the life-cycle of biogas production. Biomass Bioenerg. 2006;30(3):254–66.

[4] Buswell AM. Anaerobic Fermentators. Urbana, Illinoise: Division of the State Water Survey, State of Illinois; 1936.

[5] Symons GE, Buswell AM. The methane fermentation of carbohydrates. J Am Chem Soc. 1933;55(5):2028–36.

[6] Page DI, Hickey KL, Narula R, Main AL, Grimberg SJ. Modeling anaerobic digestion of dairy manure using the IWA Anaerobic Digestion Model No. 1 (ADM1). Water Sci Technol. 2008;58(3):689–95.

[7] Linke B. A model for anaerobic digestion of animal waste slurries. Environ Technol. 1997;18(8):849–54.

[8] Menten F, Chèze B, Patouillard L, Bouvart F. A review of LCA greenhouse gas emissions results for advanced biofuels: The use of meta-regression analysis. Renewable Sustainable Energy Rev. 2013;26:108–34.

[9] Farrell AE, Plevin RJ, Turner BT, Jones AD, O'Hare M, Kammen DM. Ethanol can contribute to energy and environmental goals. Science. 2006;311(5760):506–8.

[10] Van der Meer HG. Optimising manure management for GHG outcomes. Aust J Exp Agric. 2008;48(1-2):38–45.

[11] Massa DI, Talbot G, Gilbert Y, Massé DI. On farm biogas production: A method to reduce GHG emissions and develop more sustainable livestock operations. Anim Feed Sci Technol. 2011;166:436–45.

[12] Novak SM, Fiorelli JL. Greenhouse gases and ammonia emissions from organic mixed crop-dairy systems: a critical review of mitigation options. Agron Sustainable Dev. 2010;30(2):215–36.

[13] Nelson MC, Morrison M, Yu Z. A meta-analysis of the microbial diversity observed in anaerobic digesters. Bioresour Technol. 2011;102(4):3730–9.

[14] Miranda ND, Tuomisto HL, McCulloch MD. Meta-analysis of GHG emissions from anaerobic digestion processes in dairy farms. Environ Sci Technol. 2015;49(8):5211–9.

[15] Intergovernmental Panel of Climate Change (IPCC). Guidelines for National Greenhouse Inventories. Volume 4: Agriculture, Forestry and Other Land Use. Hayama: IPCC National Greenhuse Gas Inventories Programme; 2006.

[16] ISI Web of Knowledge [homepage on the Internet]. Search Engine Online Resource. Thomson Reuters; [cited 2015 Feb 18] Available from: http://www.isiwebofknowlege.com/.

[17] Deng Y, Xu J, Liu Y, Mancl K. Biogas as a sustainable energy source in China: Regional development strategy application and decision making. Renew Sustain Energy Rev. 2014;35:294–303.

[18] Rao PV, Baral SS, Dey R, Mutnuri S. Biogas generation potential by anaerobic digestion for sustainable energy development in India. Renew Sustain Energy Rev. 2010;14(7):2086–94.

[19] Shapiro SS, Wilk MB. An analysis of variance test for normality (complete samples). Biometrika. 1965;52(3/4):591–611.

[20] Wilcoxon F. Individual Comparisons by Ranking Methods. Biometrics Bull. 1945;1(6):80–3.

[21] Spearman C. The proof and measurement of association between two things. Am J Psychol. 1904;15(1):72–101.

[22] Brambilla M, Araldi F, Marchesi M, Bertazzoni B, Zagni M, Navarotto P. Monitoring of the startup phase of one continuous anaerobic digester at pilot scale level. Biomass Bioenerg. 2012;36:439–46.

[23] Henderson P, Seaby R. A Practical Handbook for Multivariate Methods. Pennington, UK: Pisces Conservation; 2008.

[24] Gan G, Ma C, Wu J. Data Clustering: Theory, Algorithms, and Applications. 1st ed. Philadelphia: Society for Industrial and Applied Mathematics; 2007.

[25] Lo KV, Liao PH, Bulley NR, Chieng ST. A comparison of biogas production from dairy manure filtrate using conventional and fixed-film reactors. Can Agric Eng. 1984;26(1):73–8.

[26] Lo KV, Liao PH. Psychrophilic Anaerobic Digestion of Screened Dairy Manure. Energy Agric. 1985;5(4):339–46.

[27] Lo KVV, Liao PHH, March AC. Thermophilic anaerobic digestion of screened dairy manure. Biomass. 1985;6(4):301–15.

[28] Lo KV, Liao PH, Whitehead AJ. Methane production at 22 celsius of laboratory-scale fixed-film reactors receiving screened dairy manure. Energy Agric. 1985;4(1):1–14.

[29] Lo KV, March AC, Liao PH. Determination of optimal feed strength for the thermophilic production of methane from dairy manure filtrate. Energy Agric. 1985;4(3):217–26.

[30] EPA. Final Methane to Markets International Guidance for Quantifying and Reporting the Performance of Anaerobic Digestion Systems for Livestock Manures. Washington DC: U.S. Environmental Protection Agency; 2010.

[31] Monod J. The Growth of Bacterial Cultures. Annu Rev Microbiol. 1949;3(1):371–94.

[32] Umetsu K, Takahata H, Kawamoto T. Effect of temperature on mesophilic anaerobic digestion of dairy cow slurry. Research Bulletin of Obihiro University Series I. 1992;17(4):401–8.

[33] Kaparaju PLN, Rintala JA. Effects of temperature on post-methanation of digested dairy cow manure in a farm-scale biogas production system. Environ Technol. 2003;24(10):1315–21.

[34] Belle AJ, Lansing S, Mulbry W, Weil RR. Anaerobic co-digestion of forage radish and dairy manure in complete mix digesters. Biomass Bioenerg. 2015;178:230–7.

[35] Zeeman G, Sutter K, Vens T, Koster M, Wellinger A. Psychorophilic digestion of dairy cattle and pig manure: Start-up procedures of batch, fed-batch and CSTR-type digesters. Biol Wastes. 1988;26(1):15–31.

[36] Angelidaki I, Ahring BK. Thermophilic anaerobic digestion of livestock waste: the effect of ammonia. Appl Microbiol Biotechnol. 1993;38(4):560–4.

[37] Karkkainen I, Franti P. Minimization of the Value of Davies-Bouldin Index. In: Hamza M, editor. International Conference of Signal Processing and Communication; Sep 19-22 2000; Malaga, Spain. Anaheim, CA; Calgary (AB): IASTED; 2000. p. 426–32.

[38] Pandey PK, Ndegwa PM, Soupir ML, Alldredge JR, Pitts MJ. Efficacies of inocula on the startup of anaerobic reactors treating dairy manure under stirred and unstirred conditions. Biomass Bioenerg. 2011;35(7):2705–20.

[39] Moller HB, Sommer SG, Ahring B. Methane productivity of manure, straw and solid fractions of manure. Biomass Bioenerg. 2004;26(5):485–95.

[40] Pabon-Pereira CP, De Vries JW, Slingerland MA, Zeeman G, van Lier JB. Impact of crop-manure ratios on energy production and fertilizing characteristics of liquid and solid digestate during codigestion. Environ Technol. 2014;35(19):2427–34.

[41] Demirer GN, Chen S. Effect of retention time and organic loading rate on anaerobic acidification and biogasification of dairy manure. J Chem Technol Biotechnol. 2004;79(12):1381–7.

[42] Rico C, Garcia H, Rico JL. Physical-anaerobic-chemical process for treatment of dairy cattle manure". Bioresour Technol. 2011;102(3):2143–50.

[43] Safley LM, Westerman PW. Performance of a low-temperature lagoon digester. Bioresour Technol. 1992;41(2):167–75.

[44] Park JBK, Craggs RJ. Biogas production from anaerobic waste stabilisation ponds

treating dairy and piggery wastewater in New Zealand. Water Sci Technol. 2007;55(11):257–64.

[45] Craggs R, Park J, Heubeck S. Methane emissions from anaerobic ponds on a piggery and a dairy farm in New Zealand. Aust J Exp Agric. 2008;48(1-2):142–6.

[46] Sterling MC, Lacey RE, Engler CR, Ricke SC. Effects of ammonia nitrogen on H2 and CH4 production during anaerobic digestion of dairy cattle manure. Bioresour Technol. 2001;77(1):9–18.

[47] Owen WF, Stuckey DC, Healy JB, Young LY, McCarty PL. Bioassay for monitoring biochemical methane potential and anaerobic toxicity. Water Res. 1979;13(6):485–92.

[48] Li J, Zicari SM, Cui Z, Zhang R. Processing anaerobic sludge for extended storage as anaerobic digester inoculum. Bioresour Technol. 2014 aug;166:201–10.

[49] Martí-Herrero J, Flores T, Alvarez R, Perez D. How to report biogas production when monitoring small-scale digesters in field. Biomass and Bioenergy. 2016;84:31–6.

[50] Morris GR. Anaerobic fermentation of animal wastes: A kinetic and empirical design evaluation [Thesis]. New York: Cornell University; 1976.

[51] Safley LM, Casada ME, Woodbury JW, Roos KF. Global Methane Emissions From Livestock and Poultry Manure. Washington: U.S. Environmental Protection Agency, Air and Radiation; 1992.

[52] Cumming G. Understanding The New Statistics Effect Sizes, Confidence Intervals, and Meta-Analysis. 1st ed. New York: Routledge Taylor & Francis Group; 2012.

Table Legends

Figure Captions

Tables and Figures

Table 1: Number of cases (fraction of cases[a]) reporting secondary quantitative variables according to digester type.

Secondary Variables	[Units]	Batch	Continuous	Semi-continuous
Operation conditions				
Time[b]	$[d]$	1,260 (97.0%)	326 (97.3%)	332 (64.7%)
Working volume	$[m^3]$	1,286 (99.0%)	311 (92.8%)	513 (100%)
Temperature	$[^\circ C]$	1,288 (99.1%)	291 (86.9%)	509 (99.2%)
Organic loading rate (OLR)	$[kg \cdot m^{-3}s^{-1}]$	n.a.	277 (82.7%)	377 (73.5%)
Chemical composition				
pH	n.a.	785 (60.4%)	115 (34.3%)	348 (67.8%)
Dry matter (DM)	$[kg \cdot kg^{-1}]$	1,078 (83.0%)	232 (69.2%)	470 (91.6%)
Volatile solids (VS)	$[kg \cdot kg^{-1}]$	1,063 (81.8%)	231 (69.0%)	505 (98.4%)
Chemical oxygen demand (COD)	$[kg \cdot m^{-3}]$	653 (50.3%)	183 (54.6%)	304 (59.3%)
Total Kjeldahl nitrogen (TKN)	$[kg \cdot m^{-3}]$	726 (55.9%)	159 (77.3%)	303 (59.1%)

[a] Fraction of cases that report secondary variable with respect to total cases of specific digester (in %) for that type of digester.[b]Time for batch and hydraulic retention time (HRT) for continuous and semi-continuous digesters, respectively.

Table 2: Calculations for two meta-analysis models [52].

Steps for calculation.	Fixed-Effect	Random-Effect
1. Variance of case "i"	$V_i = s_i^2/N_i$	
2. Initial weight of case "i"	$W_i = 1/V_i$	
2. Measure of heterogeneity	n.a.	$Q = \sum W_i M_i^2 - \frac{(\sum W_i M_i)^2}{\sum W_i}$
3. Std. deviation of sample	n.a	$T^2 = \frac{Q-(k-1)}{W_i - \sum W_i^2/\sum Wi}$
4. Variance of case "i"[1]	V_i	$V_i^* = V_i + T^2$
5. Weight of case "i"[1]	W_i	$W_i^* = 1/V_i^*$
6. Mean of sample	$M = \sum W_i M_i / \sum W_i$	$M^* = \sum W_i^* M_i / \sum W_i^*$
7. Variance of sample	$V_M = 1/\sum W_i$	$V_M = 1/\sum W_i^*$
8. Std. deviation	$\sqrt{V_M}$	$\sqrt{V_{M^*}}$
9. Confidence interval (CI)	$\left[M \mp z_{C/100}\sqrt{V_M}\right]$	$\left[M^* \mp z_{C/100}\sqrt{V_{M^*}}\right]$

[1]For the fixed-effect model, case variances and weights are calculated in Steps 1-2.

k: number of cases; $z_{C/100}$: critical value for confidence interval of 95% (assumed 1.96 from [52]).

Table 3: Wilcoxon's P-values between default B_0 values from IPCC and B_0 from database of this work, according to region of the world.

B_0 group	Asia, Middle East & Indian Subcontinent	West & East Europe	North America	Africa
All B_0 in database	0.023 (38)	1.000 (76)	0.908 (24)	0.571 (5)
Subgroup B_0 [a]	0.033 (7)	0.916 (36)	0.474 (8)	0.400 (2)

[a]Subgroup of "All B_0" which report specific N_i and s_i, thus, can be used for meta-analysis models. Numbers of cases of B_0 in database in parenthesis.

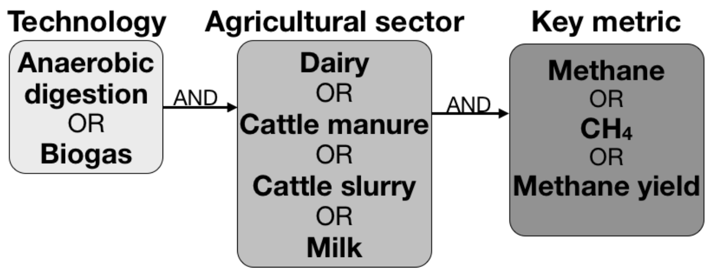

Figure 1: Keywords and logical connectors used for search of articles for the database.

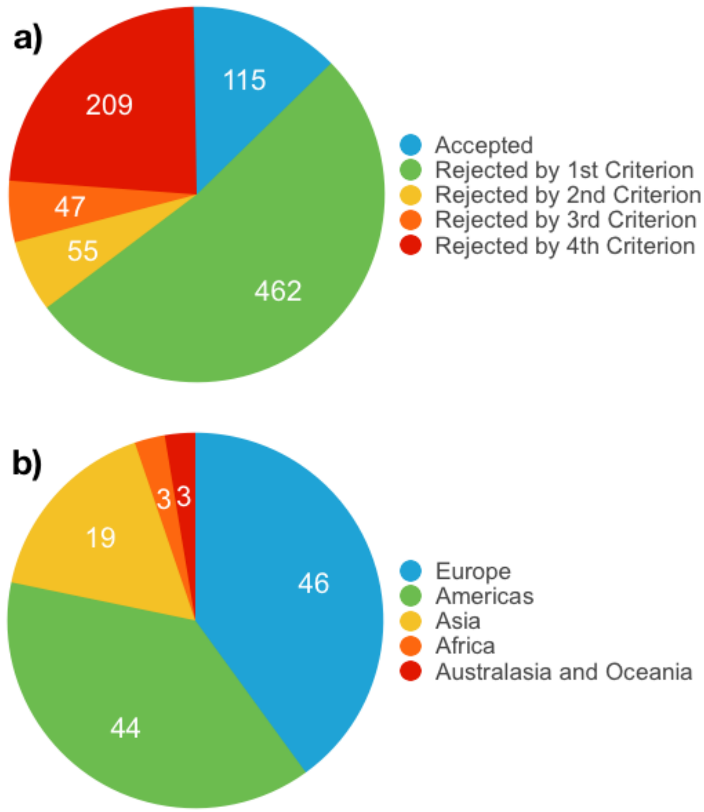

Figure 2: Pie-charts for articles in the database classified according to: a) filtering criteria and; b) continent.

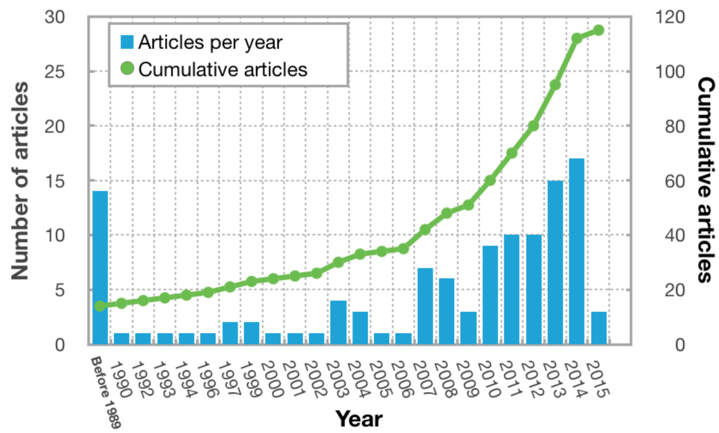

Figure 3: Articles in the database according to year of publication.

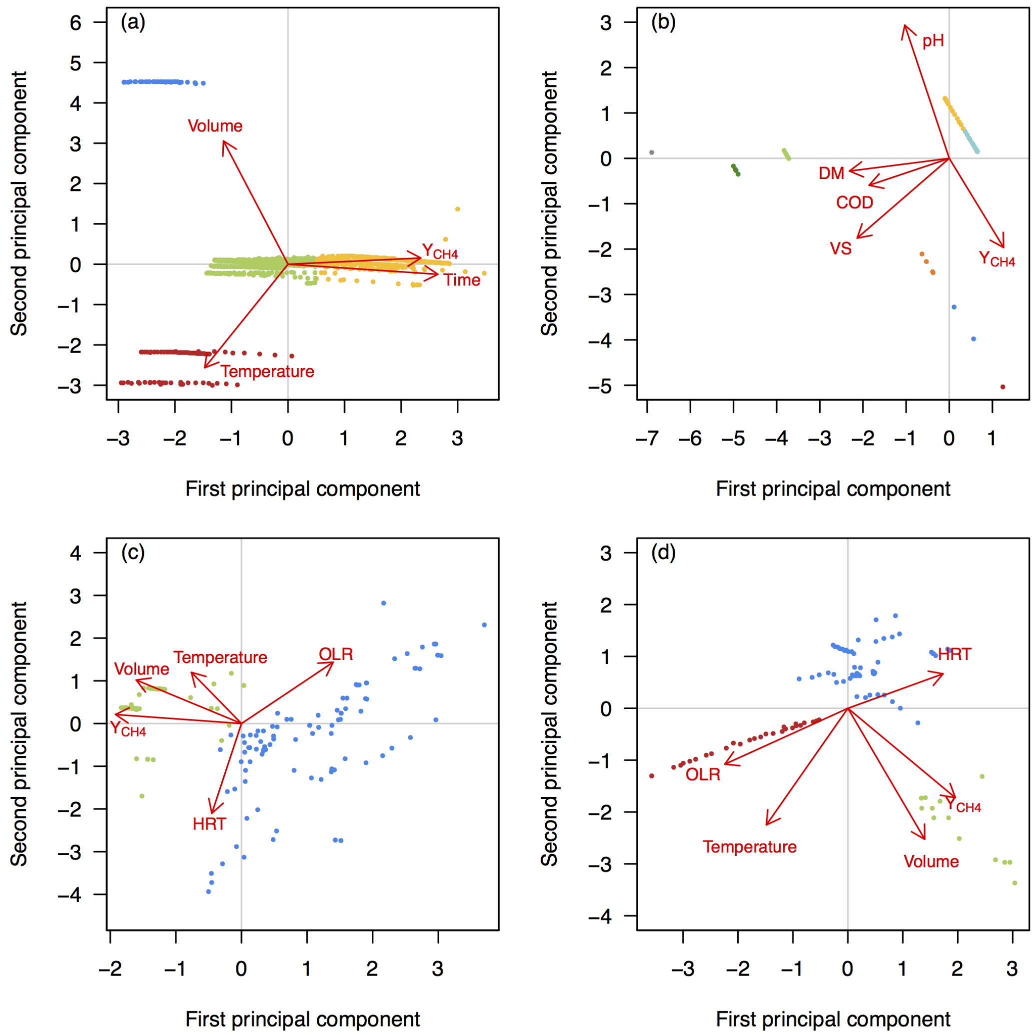

Figure 4: K-means clusters for first and second principal components of Y_{CH4} and: a) operational variables for batch; b) composition variables for batch; c) operational variables for continuous and; d) operational variables for semi-continuous digesters. Colours represent different clusters.

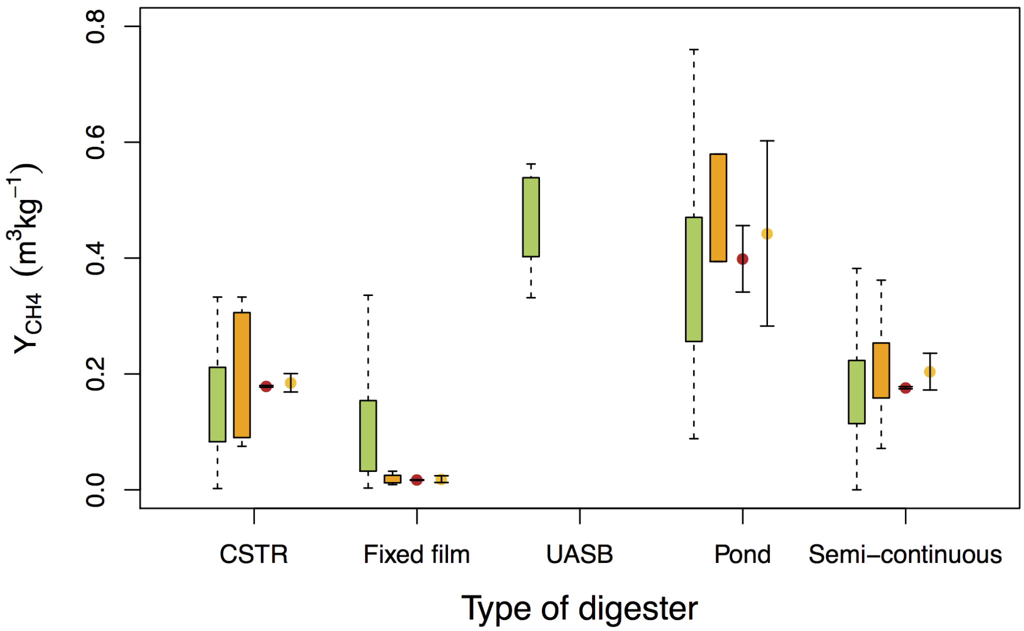

Figure 5: Y_{CH4} from different digester types found in the literature. Green box-plot: all Y_{CH4}; orange box-plot: Y_{CH4} which are reported together with s_i and N_i; red circles: results from fixed-effect mode andl; yellow circles: results from random-effect model;

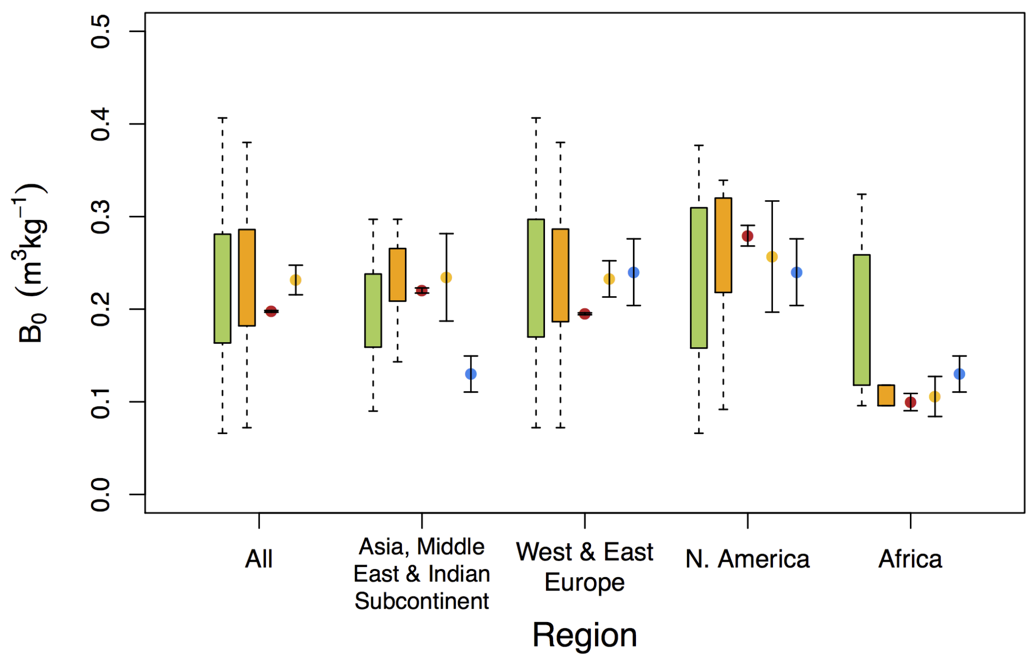

Figure 6: Maximum Y_{CH4} (B_0) from batch tests according to regions. Green box-plot: all B_0 in the developed database; orange box-plot: B_0 in database which are reported together with s_i and N_i; red circles: results from fixed-effect model; yellow circles: results from random-effect model and; blue circles: IPCC default values (IPCC 2006, [15]);