CSIRO PUBLISHING

RESEARCH PAPER
https://doi.org/10.1071/AN22082

ANIMAL PRODUCTION SCIENCE

A simplified approach for producing Tier 2 enteric-methane emission factors based on East African smallholder farm data

P. W. Ndung'u[A,B,*], C. J. L. du Toit[A], T. Takahashi[C,D], M. Robertson-Dean[E], K. Butterbach-Bahl[F], L. Merbold[G] and J. P. Goopy[A,B]

For full list of author affiliations and declarations see end of paper

*Correspondence to:
P. W. Ndung'u
Department of Animal and Wildlife Sciences, University of Pretoria, Pretoria 0002, South Africa
Email: wanjuguphyllis@ymail.com

Handling Editor:
Ermias Kebreab

Received: 3 March 2022
Accepted: 28 September 2022
Published: 28 October 2022

Cite this:
Ndung'u PW *et al.* (2022)
Animal Production Science
doi:10.1071/AN22082

ABSTRACT

Context. Accurate reporting of livestock greenhouse-gas (GHG) emissions is important in developing effective mitigation strategies, but the cost and labour requirements associated with on-farm data collection often prevent this effort in low- and middle-income countries. **Aim.** The aim of this study was to investigate the precision and accuracy of simplified activity data collection protocols in African smallholder livestock farms for country-specific enteric-methane emission factors. **Method.** Activity data such as live weight (LW), feed quality, milk yield, and milk composition were collected from 257 smallholder farms, with a total herd of 1035 heads of cattle in Nandi and Bomet counties in western Kenya. The data collection protocol was then altered by substituting the actual LW measurements with algorithm LW (ALG), feed quality (FQ) data being sourced from the Feedipedia database, reducing the need for daily milk yield records to a single seasonal milk measurement (MiY), and by using a default energy content of milk (MiE). Daily methane production (DMP) was calculated using these simplified protocols and the estimates under individual and combined protocols were compared with values derived from the published (PUBL) estimation protocol. **Key results.** Employing the algorithm LW showed good agreement in DMP, with only a small negative bias (7%) and almost no change in variance. Calculating DMP on the basis of Feedipedia FQ, by contrast, resulted in a 27% increase in variation and a 27% positive bias for DMP compared with PUBL. The substitutions of milk (MiY and MiE) showed a modest change in variance and almost no bias in DMP. **Conclusion.** It is feasible to use a simplified data collection protocol by using algorithm LW, default energy content of milk value, seasonal single milk yield data, but full sampling and analysis of feed resources is required to produce reliable Tier 2 enteric-methane emission factors. **Implications.** Reducing enteric methane emissions from the livestock is a promising pathway to reduce the effects of climate change, and, hence, the need to produce accurate emission estimates as a benchmark to measure the effectiveness of mitigation options. However, it is expensive to produce accurate emission estimates, especially in developing countries; hence, it is important and feasible to simplify on-farm data collection.

Keywords: activity data, cattle, dry-matter digestibility, GHG inventory, heart girth, milk yield, mitigation, protocol.

Introduction

In countries in sub-Saharan Africa (SSA), the contribution of agriculture to the national anthropogenic methane (CH_4) emissions may be much higher than in developed countries and can reach up to 90%, with the majority of emissions being linked to livestock production (World Bank and CIAT 2015).

Reducing enteric CH_4 emissions from ruminants is an important climate-change mitigation option, especially for countries with low levels of industrialisation (Steinfeld *et al.* 2006). Yet, for developing countries, reporting of CH_4 emissions from the livestock sector is highly uncertain due to a paucity of locally derived data on livestock production (Goopy *et al.* 2018*a*; Ndung'u *et al.* 2019; Tongwane and Moeletsi 2020). As a basis for implementing effective strategies to mitigating CH_4 emissions from the livestock sector, accurate knowledge of current emissions and a thorough understanding of underlying assumptions are essential (van Wijk *et al.* 2020). Such information is largely missing for most livestock production systems in SSA countries, and many countries are facing challenges such as inadequate knowledge of key sources of GHG emissions and missing reliable accounting systems (Merbold *et al.* 2021). This is specifically true for methane emissions from ruminants as the lack of accurate and reliable animal-activity data for local agricultural systems (Goopy *et al.* 2018*a*, 2021) hampers the development of accurate national GHG inventories. So as to close this knowledge gap, low-cost and simplified approaches to estimate GHG emissions from livestock production are the key. Thus, the main aim of this study was to evaluate simplified data collection protocols for deriving enteric CH_4-emission factors from cattle in East Africa.

The Intergovernmental Panel on Climate Change (IPCC) is the body tasked with providing guidance on GHG-emission calculations and reporting (Intergovernmental Panel on Climate Change (IPCC) 2006). It provides three frameworks for emission calculations, and these are based on the availability of data and the category of the emission source. The methodologies are in simple words split into Tier 1 (simple), Tier 2 (intermediate), and Tier 3 (complex) on the basis of application and data requirements. While Tier 1 offers global approaches that were developed on the basis of information mainly gathered in OECD countries (Goopy *et al.* 2018*a*), Tier 2 and 3 are better suited for reflecting GHG emissions from the livestock sector in specific countries, because, in these calculations, site-specific or region-specific data are used. However, Tier 1 is still the method most frequently used by developing countries because it requires only livestock census data or population estimates and a representative value of emission factor (EF) assigned to different continents and production environments (IPCC 2019). In contrast, Tier 2 and 3 require detailed characterisation of livestock, their productivity, management, and feed quality to inform livestock feed intakes otherwise known as 'activity data' that are ultimately used to estimate enteric CH_4 production (Charmley *et al.* 2016; IPCC 2019). Yet, inventories created using the Tier 1 system have been demonstrated to be considerably less representative of the actual case than are those using EFs generated through detailed investigations of enteric CH_4 emissions associated with livestock production in the Global South that take into account local conditions (e.g. seasonal fluctuations of quantity and quality of feeds, animal phenotype husbandry practices and management; du Toit *et al.* 2013; Goopy *et al.* 2018*a*, 2021).

Recently, several field studies have been undertaken in western Kenya, which collected detailed on-farm information on seasonal liveweight change, milk production, and feed composition and quality, and used this information to create spatially explicit EFs for different production systems in the region (Goopy *et al.* 2018*a*, 2021; Ndung'u *et al.* 2019, 2021). More studies using the same approach are underway by a team led by the International Livestock Research Institute for Tanzania, Uganda, and Ethiopia, so that more detailed information on CH_4 emissions from the livestock sector in East Africa will become available soon (Merbold *et al.* 2021). While these studies meet an urgent demand for better GHG-emission estimates for ruminants in SSA, the data collection is resource-intensive, requires bulky and costly equipment, and is thus lengthy and expensive to undertake. This, in turn, makes the widespread adoption of the abovementioned approach difficult to achieve. Consequently, simplified protocols, particularly those that reduce the physical and labour requirements of the sampling protocol, are urgently needed for a widespread uptake and, subsequently, the availability of spatially explicit enteric CH_4 emissions from livestock production in SSA. Thus, in this study, we investigated (1) the feasibility of simplifying the data collection protocol developed by Goopy *et al.* (2018*a*) of key activity data required in calculating Tier 2 enteric CH_4 EFs and (2) how individual and combined modifications affect the accuracy and precision of CH_4 EF estimates. We hypothesised that when replaced by a simplified protocol, some activity data are more important for accurate CH_4 emission estimates than are others. Moreover, we hypothesised that the combination of estimates of the activity data leads to larger discrepancies in CH_4 estimates than does replacing a single input variable only.

Materials and methods

Methodological approach

The approach of the previously published method by Goopy *et al.* (2018*a*) to calculate enteric CH_4 emissions from cattle in East Africa was to estimate individual animal feed intake (dry-matter intake, DMI) inferred from energy expenditure, with total energy expenditure deemed to be the sum of metabolisable energy requirement (MER) for maintenance ($MER_{Maitenance}$), growth (MER_{Growth}), milk production ($MER_{Lactation}$), and locomotion/traction ($MER_{Locomotion}$). A CH_4 yield of 20.7 g/kg DMI (Charmley *et al.* 2016) was used to calculate daily CH_4 production (DMP, g CH_4/day). For details, we refer to Goopy *et al.* (2018*a*).

Data used in this study were sourced from previously published studies conducted in Nandi and Bomet counties in western Kenya (Ndung'u *et al.* 2019, 2021). Thus, the current study investigated 257 smallholder farms with a total herd of 2270 cattle [992 females (>2 years), 103 males (>2 years), 271 heifers (1–2 years), 104 young males (1–2 years) and 800 calves (<1 year)]. However, due to the movement of animals through sales, purchases, and deaths, not all the animals were there for all four seasons and,

Table 1. Composition of the study population by animal class and region.

Class	Nandi (126 HH)	Bomet (131 HH)	Total for each class
Calves <1 year	47	13	60
Heifers 1–2 years	102	63	165
Males 1–2 years	20	21	41
Males >2 years	22	24	46
Females >2 years	358	365	723
Total	549	486	1035

hence, data used are as shown in Table 1. A set of activities including weighing cattle, heart-girth (HG) measurement, recording milk yields, milk sampling for quality analysis, feed sample collection and farm sketches should be conducted seasonally. However, these activities are time-consuming and labour-intensive and, hence, these activities were split to be undertaken after 1.5 months in each season. Therefore, smallholder farms were visited nine times within 12 months, at an interval of 1.5 months. Fig. 1 shows the activities undertaken in each farm visit. Details of farm visits, data collection, and sample analysis have been extensively reported in Goopy *et al.* (2018*a*) and Ndung'u *et al.* (2019). The following section briefly outlines the calculation of DMP and EF.

Calculation of daily CH_4 production (DMP) and emission factor (EF) by using full primary data

The DMP of cattle was determined as follows:

$$\mathrm{DMP}\ (\mathrm{CH_4\ g/day}) = \mathrm{DMI}\ (\mathrm{kg}) \times 20.7\ \mathrm{g\ CH_4/kg\ DMI} \quad (1)$$

DMI was estimated from the total MER of individual animals on a seasonal basis, calculated using algorithms derived from the CSIRO (2007). Liveweight (LW) and LW change were used to calculate MER for maintenance, growth, and locomotion, while milk yield (MY) and the energy content of milk were used to calculate MER for lactation. Feed basket (i.e. several feedstuffs available as feed for the cattle to form a feed basket) dry-matter digestibility (DMD) was used to calculate all components of the MERs (except for locomotion) and DMI. The calculated DMP for each animal for each season was then multiplied by the number of days in each season and summed across seasons to produce an annual EF (kg CH_4/head.year).

Animal LW was measured with animal weighing scales; daily MY was measured using a graduated collection vessel and recorded by farmers. To determine feed DMD, feed samples such as pasture (collected from exclusion cages), grasses grown for fodder (purchased or grown, e.g. Napier grass) and crop residues that form the feed basket were analysed for dry matter (DM), nitrogen (N), and acid detergent fibre (ADF) by using the proximate analysis method and values of N and ADF used in Eqn 2 from Oddy *et al.* (1983). Locomotion data were determined by fitting GPS collars to animals for three consecutive days to determine the distance walked per day (km).

$$\mathrm{DMD}\ (\%) = 83.58 - 0.824 \times \mathrm{ADF}\ (\%) + 2.626 \times \mathrm{N}\ (\%) \quad (2)$$

Calculation of DMP using simplified data collection protocol options

In looking to simplify and reduce resource demands of the existing published method described above (hereafter referred to as PUBL), we first reviewed the operational requirements for data collection and analysis. From this deliberation, it was concluded that LW measurements,

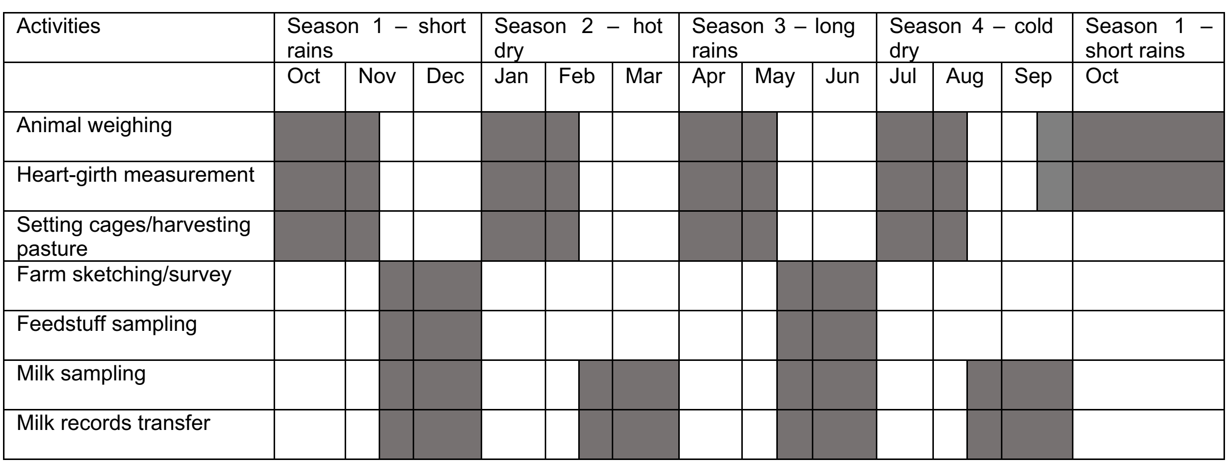

Fig. 1. Gantt chart illustrating activity data collection schedules for calculating Tier 2 enteric methane emission factors in Goopy *et al.* (2018*a*) and Ndung'u *et al.* (2019).

assessing milk production, and determination of the feed basket were activities that required high levels of resources.

Simplifying measurements of LW

Assessment of LW and LW change is a key measurement for determining energy requirements for maintenance, growth, and locomotion (CSIRO 2007). The use of animal weighing scale requires a four-wheeled pick-up truck and field assistance to facilitate setting up the scales. Substituting HG measurements for scales would represent (a) a large saving in time, (b) reduce field staff requirement, and (c) remove the necessity for heavy-duty vehicles. Measurements of HG were routinely undertaken during the original studies and were used to assess the accuracy and precision of LW estimates made using an algorithm (Eqn 3) developed for an independent population (Goopy *et al.* 2018*b*).

$$\mathrm{LW\ (kg)} = 73.599 - 2.291 \times \mathrm{HG\ (cm)} + 0.02362 \times \mathrm{HG}^2\ \mathrm{(cm)} \quad \mathrm{Adj.}R^2 = 0.856 \tag{3}$$

The dry-matter digestibility of the feed basket

Feed composition affects nutritive value, intake, and, subsequently, enteric CH_4 production. The baseline (full primary data) method assessed involved ere was based on both periodic collections of representative feed samples and their proximate analysis using wet chemistry to provide the most accurate assessment achievable of the DMD of feed baskets in different agro-ecological zones and seasons. Feed-basket DMD was used repeatedly through re-calculation based on intake variation and is, thus, highly relevant for DMP. The resources required to gather and record the information in the field is one cost, whereas chemical analysis of feed samples caused substantial costs, particularly in countries that do not have easy access to nutritional laboratories. We assessed the effect of substituting our field-measured values of DMD with the means of published values from the Feedipedia database (https://www.feedipedia.org/).

Milk quantity and quality

Energy for lactation may be the second greatest component of energy expenditure for the lactating animals (Dong *et al.* 2006). As such, data on milk quality and quantity are considered crucial for the determination of energy expenditure. In theory, MY can be fairly simply obtained using basic measuring equipment and trained farmers. However, from field studies, this often turns out to be a challenge as smallholder farmers were not regularly used to record-keeping. Furthermore, in some livestock production systems that focus on meat, not milk (e.g. extensive pastoral systems), milk is often only obtained on household demand. Due to the uncommon tendency of record-keeping, MY records were often rather uncertain. Instead of using an average of daily MY records kept by farmers, we considered spot sampling where MY measurements were requested only on the day before the researcher's farm visit.

Determining the composition of milk requires the collection, careful handling, transport, and subsequent laboratory analysis of milk samples for butterfat (BF) and solid-non-fat (SNF) content to calculate the energy contenet of milk (E) (Eqn 4); hence, this needs a milk analysis equipment such as a lactoscan. This is a significant drain on project resources. We examined the effect of substituting a single published energy content of milk value (3.054 MJ per kilogram of fat-corrected milk (FCM); CSIRO 2007) on overall estimates of milk energy.

$$E\ \mathrm{(MJ/kg\ FCM)} = 0.0386 \times \mathrm{BF\ (\%)} + 0.0205 \times \mathrm{SNF\ (\%)} - 0.0236 \tag{4}$$

Summary of simplified measurement protocols

We assessed the effect of substituting individually four data inputs that would simplify measurements and significantly reduce resource requirements for the published method. In addition, we assessed the effects of each combination of these possible substitutions on the loss in accuracy and precision.

The following abbreviations are used to describe changes to the standard protocol:

ALG: values for LW were derived from an algorithm (ALG) by using HG measurements.

FQ: the indirectly measured DMD (derived from feed-quality (FQ) values of nitrogen and acid detergent fibre) was based on values reported in the Feedipedia database, i.e. not using its measured values from the local laboratory.

MiY: this was assessed for female adult cows (>2 years) only. The average of daily milk yield (MiY) records kept by farmers were substituted with a single MY (sum of morning and evening milk) of the day before the 3-monthly farm visit for MY-recording activity in the calculations of MER for lactation.

MiE: this was assessed for female adult cows (>2 years) only. The milk energy (MiE) derived from own measurements of milk BF and SNF was substituted with a default energy content of milk value from CSIRO (2007) in the calculations of MER for lactation.

Combinations among all the altered data collection protocols were also investigated.

Data analyses

To assess how simplifications of methods affect DMP and, thus, annual EF estimates, a sensitivity analysis was conducted. The recalculated DMP and EFs from alternative data inputs were assessed independently and in combination.

DMP was calculated for individual animals for four seasons and, thereafter, an average of all the seasons was calculated using R (R Core Team 2021). There were 1035 counts of DMP, one for each animal across all four seasons in the study (Table 1). A repeat of DMP calculations was undertaken using the ALG, FQ, and ALG + FQ measurement protocols for all animals and thus also animal classes. The difference between the DMP derived from the altered measurement protocols and the PUBL DMP were calculated for each animal and the average and standard deviation (s.d.) of these were obtained to evaluate the loss of accuracy and precision in the new protocols. Twelve more calculations of DMP were run for females >2 years (n = 723) to include milk quantity-adjusted (MiY) and quality-adjusted (MiE) measurement protocols and their two-, three-, and four-way combinations with ALG and FQ data collection protocols. Similarly, the differences between the DMP for each of the altered protocols and the PUBL DMP were found, and the accuracy and precision were obtained for these.

A linear mixed-effects model was fitted to the data to verify the statistical significance of the differences for each protocol in comparison to the PUBL DMP. Animal age and Region (Nandi, Bomet) were included as variables in this model.

Results

Our calculated DMP based on the simplified data collection approach showed a substantial deviation from DMP calculated on the existing standard data collection protocol (Fig. 2). Yet, the deviation was not uniform. Replacing LW

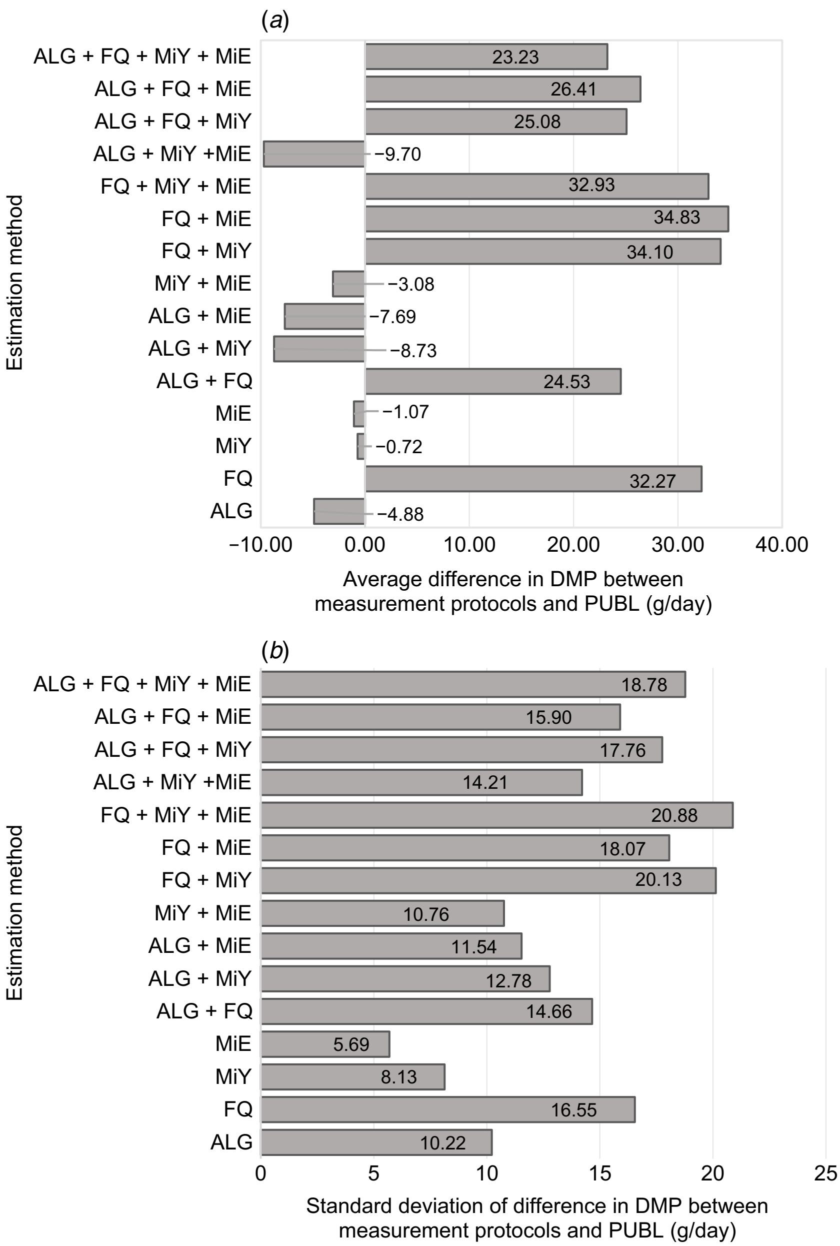

Fig. 2. (*a*) Average difference in daily methane production (g/day) for all cattle estimated in different estimation protocols of liveweight (ALG) and feed quality (FQ) and their combination, and for all female cattle for protocols involving MiY and MiE. All protocols are compared with the reference protocol PUBL for each animal, which has an average DMP of 122 g/day for all cattle and 139 g/day for females (>2 years) only. (*b*) The standard deviation of the difference in daily methane production estimated under ALG, FQ, MiY, and MiE and their combination when compared with the published (PUBL) protocol.

measurements by LW estimated through HG measurements resulted in 4.88 g/day lower DMP, whereas substituting FQ results resulted in a considerably larger DMP (32.27 g/day higher than published values). There was a substantial increase in variation on the use of algorithm LW (ALG) and Feedipedia FQ when compared to PUBL values, i.e. s.d. = 10.22 and s.d. = 6.55 respectively.

Combining two substitution measures tended to result in increased variance when compared with the measure that underestimates (ALG) and a slight reduction in variance when compared with the measure that overestimates (FQ).

Females (>2 years) have an energy requirement for lactation that is part of the total energy requirement, and it is derived using milk yield and energy content of milk. The substitutions related to lactation (MiY and MiE) showed modest variation from the original measured values (s.d. = 8.18 and 5.69) and almost no bias (−1.07 and 0.72). Again, the substitution of DMD estimates from wet chemistry with mean values from the Feedipedia database in combination with both lactation substitutions showed large deviations and a substantial bias of 34.10 (FQ and MiY) and 34.83 (FQ and MiE) g/day. In general, three-way, and four-way combinations led to a slight improvement in precision.

Table 2 presents the comparison between the PUBL DMP and the simplified protocols. A linear mixed-effects model confirmed statistical significance for the differences for every measurement protocol compared with PUBL ($P < 0.00005$) (see Table 3). However, there were no regional differences in the DMP calculated (P-value = 0.4530)

Discussions

The accuracy of estimating enteric CH_4 emissions by using models has been a key interest with numerous studies (du Toit *et al.* 2013; Ndung'u *et al.* 2019; Goopy *et al.* 2021; Ndao 2021). These studies largely agree on two conclusions, namely that (1) appropriate assumptions need to be made that reflect the conditions of the livestock system under investigation, and (2) data inputs to be used by the model need to have sufficient accuracy to derive reliable estimates of the actual emission situation. The latter has been a challenge for developing countries due to the huge capital investments needed to achieve an up-to-date database on livestock parameters needed for such models.

The accurate estimation of DMP facilitates the creation of informed, region-specific EF for ruminant livestock, which is vitally important for the development of reliable GHG inventory and, in turn, mitigation strategies for SSA countries. Although the data collection protocols for the recently completed studies (Goopy *et al.* 2018*a*; Ndung'u *et al.* 2019), which formed our reference data in this work, are robust and sound, it was recognised that the resources such as animal weighing scales, a utility vehicle, and several field assistants needed to undertake further studies may outstrip the capacity of some potential researchers. Thus, new approaches were sought to reduce the equipment and other resources, while minimising loss of veracity. Practical alternatives were found for the parameters that absorbed

Table 2. A comparison of daily methane production (g/day) derived using the published protocol (PUBL) and simplified protocols of measurements of liveweight (ALG), feed quality (FQ), milk yield (MiY) and energy content of milk (MiE) and their combinations.

Parameter/combination	Published protocol (PUBL)			Simplified protocols		
	Average DMP (g/day)	s.e.m	*n*	Average DMP (g/day)	s.e.m	*n*
ALG	121.8	1.79	1036	116.9	1.80	1036
FQ	121.8	1.79	1036	154.1	2.27	1036
MiY	138.4	2.48	723	141.5	2.56	723
MiE	137.3	2.70	723	136.0	2.61	723
ALG and FQ	121.8	1.79	1036	146.3	2.21	1036
ALG and MiY	138.8	2.48	723	135.8	2.36	723
ALG and MiE	138.8	2.48	723	132.6	2.19	723
MiY and MiE	138.8	2.48	723	140.9	2.46	723
FQ and MiY	138.8	2.48	723	178.3	3.37	723
FQ and MiE	138.8	2.48	723	172.1	3.20	723
FQ and MiY and MiE	138.8	2.48	723	177.5	3.26	723
ALG and MiY and MiE	138.8	2.48	723	135.1	2.25	723
ALG and FQ and MiY	138.8	2.48	723	169.5	2.98	723
ALG and FQ and MiE	138.8	2.48	723	165.7	2.79	723
AlG and FQ and MiY and MiE	138.8	2.48	723	167.3	2.82	723

Simplified protocols are ALG, algorithm weight; FQ, Feedipedia feed-quality values; MiY, single milk yield measurement per season; MiE, default energy content of milk.

Table 3. The difference in daily methane production (DMP, g/day) between the published protocol and simplified data collection protocols and the significance levels (*P*-value) when compared with published values.

Variable	Difference	*P*-value
(intercept)	61.1	0.0000***
Protocol		
ALG	−4.9	0.000***
FQ	32.3	0.000***
MiY	−2.5	0.000***
MiE	−2.8	0.000***
ALG + FQ	24.5	0.000***
ALG + MiY	−10.5	0.000***
ALG + MiE	−9.4	0.000***
MiY + MiE	−4.8	0.000***
FQ + MiY	32.4	0.000***
FQ + MiE	33.1	0.000***
FQ + MiY + MiE	31.3	0.000***
ALG + MiY + MiE	−11.4	0.000***
ALG + FQ + MiY	23.3	0.000***
ALG + FQ + MiE	24.7	0.000***
AlG + FQ + MiY + MiE	21.5	0.000***
Region		
Nandi	1.9	0.4530[n.s.]

The DMP derived by simplified protocols in the table is compared with the published DMP (PUBL). Nandi region is compared with Bomet region.
****P* < 0.01. n.s., not significant.
Simplified protocols are: ALG, algorithm weight; FQ, Feedipedia feed-quality values; MiY, single milk yield measurement per season; MiE, default energy content of milk.

the greater amount of resources, namely LW, milk analysis and yield, and feed characteristics, and these were compared using estimates of DMP for the substitution, with our previously published estimates.

Heart-girth algorithm to estimate LW

Animals' LW and LW change are the most pervasive and arguably the most influential measurements used in the calculation of DMP and, thereby, EFs. It is integral to the calculation of $MER_{Maitenance}$, MER_{Growth}, and of $MER_{Locomotion}$ and, so, the importance of the accurate measurement of LW is difficult to overstate. The results indicated that while ALG has a similar degree of precision to the use of animal scales, the measurement exhibits a degree of systematic bias, resulting in an underestimation of DMP by 7% (resulting in an EF reduction of ~2 kg CH_4/head.year for an average-sized animal). This is not ideal, and a possible solution is to employ a better algorithm. The algorithm used in this study was developed from cattle populations from West and East Africa, comprising 1513 cattle. The original study (Goopy *et al.* 2018*b*) was found to have very good agreement with gravimetric measurement (Adj.R^2 = 0.920) and a prediction error similar to that of the aggregated dataset for African cattle that could be found in the literature (see Table 4). Although the weight of animals in this study significantly differed from that of the the dataset used for establishing the HG relationship, the prediction error remained similar to that in the original study. This suggests that the algorithm used is unlikely to be able to be further improved. If this is the case, the implications of the entrenched bias must be weighed against the tactical advantage of a simple HG measurement over cumbersome and costly scales.

It must also be emphasised that algorithms developed within a particular population may be less accurate when employed outside of the population, especially if the algorithm is based on a simple linear regression model (Goopy *et al.* 2018*b*). Therefore, if an HG algorithm is to be used to assess LW in a different population, it is strongly recommended that the decision be informed by a preliminary assessment of the model's predictive capacity among the population where it is to be employed.

Feed quality – assessment of DMD

Dry-matter digestibility is an important input in the estimation of DMP, being a key determinant of intake (for a given energy requirement), and used in the calculation of $MER_{Maintenance}$, $MER_{Lactation}$, and DMI. Determination of DMD is an elaborate process involving several chemical assays, which is often difficult to perform in research laboratories in SSA countries, so it was considered that substituting direct analysis for values freely available from reputable literature (i.e. Feedipedia) could confer significant advantages. The use of values from Feedipedia resulted in an overestimation of DMP by 24%. The causes underlying this are clear, while the magnitude of the error was not hypothesised. Studies such as those of Ndung'u *et al.* (2019) have demonstrated the spatial and temporal variability in the composition of livestock feed, which our sampling during the establishment of the database used in this study was able to capture. By contrast, Feedipedia values do typically cover a range, we were simply limited to using a mid-range value in all calculations. Given the large reductions in both precision and accuracy, it is not recommended that this measurement change be adopted until region-specific feed information is available in Feedipedia.

Lactation parameters

The substituted measures for MiE and MiY had the least impact on the DMP among the measures explored, with very little loss of accuracy (−0.1%) or precision (−0.02%). Concerning MiE, this was clearly because the value adopted from CSIRO (2007) was well representative of the actual

Table 4. Comparison of prediction errors at 75th, 90th, and 95th percentiles for Quadratic model to estimate liveweight by using heart-girth measurements between datasets from Goopy *et al.* (2018*b*) study, and the present study.

Study site	Study	Prediction error (percentiles)		
		75th (%)	90th (%)	95th (%)
East Africa dataset	Goopy *et al.* (2018*b*)	±15	±23	±30
West Africa dataset	Goopy *et al.* (2018*b*)	±14	±21	±27
Aggregated dataset	Goopy *et al.* (2018*b*)	±15	±22	±29
Validation dataset-Nyando	Goopy *et al.* (2018*b*)	±10	±15	±18
Nandi dataset	Present study	±19	±25	±29
Bomet dataset	Present study	±13	±19	±22

case in the populations studied. This was surprising, given that the value was derived from European cattle, although the work of Cheruiyot *et al.* (2018) suggested that variation in milk composition across commonly encountered African phenotypes is minimal because the milk samples were from a sample pool with dominant Holstein–Friesian origin, which is commonly found in the European farms. However, the latter statement remains disputed as a study by Kebede (2018) brought this assertion into question due to the breed effect on milk fat. In view of this, it is suggested that adopting the given value for milk energy content is likely a 'no-regrets' option for further studies of this kind, but it is recommended that the milk from a small representative subsample be analysed for comparison before committing to fieldwork.

Our data showed that a single day's milk collection (MiY) provided nearly identical estimates of seasonal milk production, as compared to the daily collection throughout the season, and that decreases in accuracy remained 0.03% lower and 0.01% higher in variation. This is likely the case for two important and inter-related reasons. First, the lactation curves of smallholder cattle have been characterised as being of a slow, steady decline (Lobago *et al.* 2007), meaning that there is little variation in production from the beginning to the end of lactation. The second reason is that milk production itself is typically so low in these systems (<5 L/day for mixed crop–livestock systems and <2 L/day for pastoral and agropastoral systems; Garnsworthy *et al.* 2012; Ndung'u *et al.* 2019), making the energy required for milk production less important to the overall calculation of DMP. While our results indicated that only single-day data are required to determine seasonal yield, there may be considerable risks associated with adopting such a practice. The dichotomous nature of high and low yield dairy systems in South Africa (Abin *et al.* 2018) highlights the association among high production, good nutrition, and pronounced lactation curves in cattle (and *vice versa* for low-production systems). This suggests that if a livestock production system is understood and production is known to be low, the 'single sample' protocol may be adopted. Otherwise, repeated sampling is indicated, not least of all because, as production increases, the importance of $MER_{Lactation}$ to overall energy expenditure also rises.

Operation costs

To compare the effect of simplifying the data collection protocols to the cost of operations, the relative contribution of transport logistics, expertise (staff), and equipment were computed on the basis of Kenya's case-study projects costs incurred, i.e. 30%, 28%, and 20% of the total budget respectively. Transport costs would be reduced significantly when the need for a heavy track was avoided as the fuel consumption would be reduced by half while the cost of hiring the heavy vehicle was reduced as well. Replacing daily milk measurements with spot sampling would completely remove the cost of providing measuring cans. Similarly, the cost of Lactoscan would be removed as well when the use

Table 5. Sources of expenditure and the comparison of their relative cost contribution to the overall budget for data collection for the published protocol (PUBL) with simplified data collection protocols that use algorithm liveweight (ALG), single milk yield measurement (MiY), and default milk energy (MiE) and their combinations.

Protocol	Relative cost contribution to overall budget (%)				Total budget
	Transport (vehicle + fuel + driver)	Expertise (staff)	Equipment (animal scale + measuring cans + Lactoscan)	Laboratory (feed analysis)	
PUBL	30.0	28.0	20.0	22.0	100.0
ALG	20.0	27.0	13.0	22.0	82.0
MiY	30.0	28.0	9.0	22.0	89.0
MiE	30.0	28.0	17.0	22.0	97.0
ALG + MiY	20.0	27.0	2.5	22.0	71.5
ALG + MiE	20.0	27.0	10.8	22.0	79.8
ALG + MiY + MiE	20.0	27.0	0.2	22.0	69.2

of default milk energy value were adopted. The cost of nutrition laboratory services (22% of the total budget) was not altered, as changing the source of the feed-quality data showed a high loss of accuracy, as discussed earlier in DMP estimates and, hence, not recommended. The combination of the use of algorithm LW (ALG), single MY measurement (MiY) and default milk energy (MiE) would cause an overall budget cost reduction of ~30%, as shown in Table 5.

Conclusions

This study has led us to the following key conclusions:

1. Actual measurements of LW data can be substituted with modelled LW data by using the HG algorithm. This still produces DMP estimates with a low bias and variation.
2. Feed-quality data cannot be substituted by averaged literature data as it causes a wide variation and highly biased results. Hence, diet-quality information should not be compromised at any point and should be accurately determined to have better CH_4 estimates.
3. Using the default energy content of milk representative of local dairy instead of actual measurements of milk-quality contents only caused a bias of –1.07 g/day and, hence, can be adopted.
4. In the context of a low-producing dairy production system, a single milk yield recorded the day before a farm visit per season instead of a seasonal average for calculating energy requirements for lactation only caused a bias of 0.72 g/day and, hence, can be adopted.
5. We can have a simplified protocol by using algorithm LW, default milk energy, and a single milk yield data and measured feed quality (ALG + MiY + MiE) that would produce DMP estimates with an error in DMP of ≤9.7 g/day, with reduced operation cost of ~30% fo the total budget.

References

Abin S, Visser C, Banga CB (2018) Comparative performance of dairy cows in low-input smallholder and high-input production systems in South Africa. *Tropical Animal Health and Production* **50**, 1479–1484. doi:10.1007/s11250-018-1584-9

Charmley E, Williams SRO, Moate PJ, Hegarty RS, Herd RM, Oddy VH, Reyenga P, Staunton KM, Anderson A, Hannah MC (2016) A universal equation to predict methane production of forage-fed cattle in Australia. *Animal Production Science* **56**, 169–180. doi:10.1071/AN15365

Cheruiyot EK, Bett RC, Amimo JO, Mujibi FDN (2018) Milk composition for admixed dairy cattle in Tanzania. *Frontiers in Genetics* **9**, 142. doi:10.3389/fgene.2018.00142

CSIRO (2007) 'Nutrient requirements for domesticated ruminants.' (Eds HDM Freer, H Dove, JV Nolan) (CSIRO Publishing: Melbourne, Vic., Australia)

Dong H, Mangino J, McAllister TA, Hatfield JL, Johnson DE, Lassey KR, Aparecida de Lima M, Romanovskaya A (2006) Emissions from livestock and manure management. In 'Agriculture, forestry and other land use, IPCC Guidelines for National Greenhouse Gas Inventories'. (Eds HS Eggelston, L Buendia, K Miwa, T Ngara, K Tanabe) pp. 87. (Institute for Global Environmental Strategies (IGES): Hayama, Japan)

du Toit CJL, Meissner HH, van Niekerk WA (2013) Direct methane and nitrous oxide emissions of South African dairy and beef cattle. *South African Journal of Animal Science* **43**, 320–339. doi:10.4314/sajas.v43i3.7

Garnsworthy PC, Craigon J, Hernandez-Medrano JH, Saunders N (2012) Variation among individual dairy cows in methane measurements made on farm during milking. *Journal of Dairy Science* **95**, 3181–3189. doi:10.3168/jds.2011-4606

Goopy JP, Onyango AA, Dickhoefer U, Butterbach-Bahl K (2018*a*) A new approach for improving emission factors for enteric methane emissions of cattle in smallholder systems of East Africa – results for Nyando, western Kenya. *Agricultural Systems* **161**, 72–80. doi:10.1016/j.agsy.2017.12.004

Goopy JP, Pelster DE, Onyango A, Marshall K, Lukuyu M (2018*b*) Simple and robust algorithms to estimate liveweight in African smallholder cattle. *Animal Production Science* **58**, 1758–1765. doi:10.1071/AN16577

Goopy JP, Ndung'u PW, Onyango A, Kirui P, Butterbach-Bahl K (2021) Calculation of new enteric methane emission factors for small ruminants in western Kenya highlights the heterogeneity of smallholder production systems. *Animal Production Science* **61**, 602–612. doi:10.1071/AN19631

Intergovernmental Panel on Climate Change (IPCC) (2006) 'IPCC Guidelines for National Greenhouse Gas Inventories. Prepared by National Greenhouse Gas Inventory Programme' (Eds HS Eggleston, L Buendia, K Miwa, T Ngara, K Tanabe) (Vol. 1: General Guidance and Reporting) (Institute for Global Environmental Strategies (IGES): Hayama, Japan)

IPCC (2019) Chapter 10: Emissions from livestock and manure management. In '2019 Refinement to the 2006 Guidelines for National Greenhouse Gas Inventories. Vol. 4'. Agriculture, Forestry and Other Land Use. (Eds O Gavrilova, A Leip, H Dong, JD MacDonald, CA Gomez Bravo, B Amon, R Barahona Rosales, Ad Prado, MAd Lima, W Oyhantçabal, TJvd Weerden, Y Widiawati) (IPCC)

Kebede E (2018) Effect of cattle breed on milk composition in the same management conditions. *Ethiopian Journal of Agricultural Sciences* **28**, 53–63.

Lobago F, Bekana M, Gustafsson H, Kindahl H (2007) Longitudinal observation on reproductive and lactation performances of smallholder crossbred dairy cattle in Fitche, Oromia region, central Ethiopia. *Tropical Animal Health and Production* **39**, 395–403. doi:10.1007/s11250-007-9027-z

Merbold L, Scholes RJ, Acosta M, Beck J, Bombelli A, Fiedler B, Grieco E, Helmschrot J, Hugo W, Kasurinen V, Kim D-G, Körtzinger A, Leitner S, López-Ballesteros A, Ndisi M, Nickless A, Salmon E, Saunders M, Skjelvan I, Vermeulen AT, Kutsch WL (2021) Opportunities for an African greenhouse gas observation system. *Regional Environmental Change* **21**, 104. doi:10.1007/s10113-021-01823-w

Ndao S (2021) Analysis of inputs parameters used to estimate enteric methane emission factors applying a Tier 2 model: case study of native cattle in Senegal. In 'Animal feed science and nutrition: health and environment'. (IntechOpen) doi:10.5772/intechopen.99810

Ndung'u PW, Bebe BO, Ondiek JO, Butterbach-Bahl K, Merbold L, Goopy JP (2019) Improved region-specific emission factors for enteric methane emissions from cattle in smallholder mixed crop: livestock systems of Nandi County, Kenya. *Animal Production Science* **59**, 1136–1146. doi:10.1071/AN17809

Ndung'u PW, Kirui P, Takahashi T, du Toit CJL, Merbold L, Goopy JP (2021) Data describing cattle performance and feed characteristics to calculate enteric methane emissions in smallholder livestock systems in Bomet County, Kenya. *Data in Brief* **39**, 107673. doi:10.1016/j.dib.2021.107673

Oddy V, Robards G, Low S (1983) Prediction of *in vivo* dry matter digestibility from the fibre and nitrogen content of a feed. In 'Feed information and animal production: proceedings of the second symposium of the International Network of Feed Information Centres'. (Eds GE Robards, RG Packham) (Commonwealth Agricultural Bureaux: Farnham Royal, Slough [Buckingham], UK)

R Core Team (2021) R: A language and environment for statistical computing. R Foundation for Statistical Computing. Available at https://www.R-project.org/

Steinfeld H, Wassenaar T, Jutzi S (2006) Livestock production systems in developing countries: status, drivers and trends. *Revue Scientifique et Technique* **25**, 505–516. doi:10.20506/rst.25.2.1677

Tongwane MI, Moeletsi ME (2020) Emission factors and carbon emissions of methane from enteric fermentation of cattle produced under different management systems in South Africa. *Journal of Cleaner Production* **265**, 121931. doi:10.1016/j.jclepro.2020.121931

van Wijk MT, Merbold L, Hammond J, Butterbach-Bahl K (2020) Improving assessments of the three pillars of climate smart agriculture: current achievements and ideas for the future. *Frontiers in Sustainable Food Systems* **4**, 558483. doi:10.3389/fsufs.2020.558483

World Bank and CIAT (2015) 'Climate-smart agriculture in Kenya.' CSA Country Profiles for Africa, Asia, and Latin America and the Caribbean Series. (The World Bank Group: Washington, DC, USA) Available at https://hdl.handle.net/10568/69545. [Accessed 7 November 2017]

Data availability. The data that support this study are available in Mendeley Data (https://doi.org/10.17632/j5b9d7dd2b.2) and Ndung'u *et al.* (2019) (https://doi.org/10.1071/AN17809 and Supplementary material https://www.publish.csiro.au/an/acc/AN17809/AN17809_AC.pdf).

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

Declaration of funding. This study was funded by the International Fund for Agricultural Development (IFAD) under the 'Greening Livestock: Incentive-based Interventions for Reducing the Climate Impact of Livestock in East Africa' project (Grant No. 2000000994). The research was further supported by the German Federal Ministry for Economic Cooperation and Development (BMZ issued through GIZ) Programme of Climate Smart Livestock (PCSL). The authors appreciate the support from the South African National Research Foundation, NRF Thuthuka Grant No. TTK180419322838.

Acknowledgements. This study became possible only due to the availability of the analytical capacity within the Mazingira Centre, and Sustainable Livestock Systems (SLS) program at International Livestock Research Institute (ILRI), Nairobi, Kenya. The authors acknowledge the CGIAR Fund Council, Australia (ACIAR), Irish Aid, European Union, IFAD, the Netherlands, New Zealand, UK, USAID, and Thailand for funding the CGIAR Research Program on Livestock.

Author affiliations

[A]Department of Animal and Wildlife Sciences, University of Pretoria, Pretoria 0002, South Africa.

[B]Mazingira Centre, International Livestock Research Institute, PO Box 30709-00100, Nairobi, Kenya.

[C]Rothamsted Research, North Wyke, Okehampton, Devon EX20 2SB, UK.

[D]University of Bristol, Langford House, Langford, Somerset BS40 5DU, UK.

[E]School of Mathematics, University of New England, Armidale, NSW, Australia.

[F]Institute for Meteorology and Climate Research, Atmospheric Environmental Research, Karlsruhe Institute of Technology, D-82467 Garmisch-Partenkirchen, Germany.

[G]Agroscope, Research Division Agroecology and Environment, Reckenholzstrasse 191, 8046 Zurich, Switzerland.