

Benchmarking the environmental performance of specialized milk production systems: selection of a set of indicators



W. Mu^{a,b,*}, C.E. van Middelaar^a, J.M. Bloemhof^b, B. Engel^c, I.J.M. de Boer^a

^a Animal Production Systems group, Wageningen University, P.O. Box 338, 6700 AH Wageningen, the Netherlands

^b Operations Research and Logistics group, Wageningen University, P.O. Box 8130, 6700 EW Wageningen, the Netherlands

^c Biometris, Wageningen University, P.O. Box 16, 6700 AA Wageningen, the Netherlands

ARTICLE INFO

Article history:

Received 14 March 2016

Received in revised form 1 July 2016

Accepted 6 August 2016

Available online 12 August 2016

Keywords:

Life cycle assessment

Nutrient balance

Livestock

Sustainability

Environmental indicators

Correlation analysis

ABSTRACT

Dairy production across the world contributes to environmental impacts such as eutrophication, acidification, loss of biodiversity, and use of resources, such as land, fossil energy and water. Benchmarking the environmental performance of farms can help to reduce these environmental impacts and improve resource use efficiency. Indicators to quantify and benchmark environmental performances are generally derived from a nutrient balance (NB) or a life cycle assessment (LCA). An NB is relatively easy to quantify, whereas an LCA provides more detailed insight into the type of losses and associated environmental impacts. In this study, we explored correlations between NB and LCA indicators, in order to identify an effective set of indicators that can be used as a proxy for benchmarking the environmental performance of dairy farms. We selected 55 specialised dairy farms from western European countries and determined their environmental performance based on eight commonly used NB and LCA indicators from cradle-to-farm gate. Indicators included N surplus, P surplus, land use, fossil energy use, global warming potential (GWP), acidification potential (AP), freshwater eutrophication potential (FEP) and marine eutrophication potential (MEP) for 2010. All indicators are expressed per kg of fat-and-protein-corrected milk. Pearson and Spearman Rho's correlation analyses were performed to determine the correlations between the indicators. Subsequently, multiple regression and canonical correlation analyses were performed to select the set of indicators to be used as a proxy. Results show that the set of selected indicator, including N surplus, P surplus, energy use and land use, is strongly correlated with the eliminated set of indicators, including FEP ($r = 0.95$), MEP ($r = 0.91$), GWP ($r = 0.83$) and AP ($r = 0.79$). The canonical correlation between the two sets is high as well ($r = 0.97$). Therefore, N surplus, P surplus, energy use and land use can be used as a proxy to benchmark the environmental performance of dairy farms, also representing GWP, AP, FEP and MEP. The set of selected indicators can be monitored and collected in a time and cost-effective way, and can be interpreted easily by decision makers. Other important environmental impacts, such as biodiversity and water use, however, should not be overlooked.

© 2016 Elsevier Ltd. All rights reserved.

1. Introduction

Dairy products are important protein sources in human diets. Around 57% of the protein content of an average European diet consists of livestock products, of which about one third is milk-derived (FAOSTAT, 2013). The global demand for milk products is expected to increase further due to population growth, rising incomes and on-going urbanization (FAO, 2006).

Dairy production, however, has a major impact on the environment. The global dairy sector, producing both milk and meat, for example, is responsible for about 30% of the anthropogenic greenhouse gas (GHG) emissions from livestock (Gerber et al., 2013). Dairy production across the world is shown to contribute also to eutrophication, acidification, loss of biodiversity, and use of resources, such as land, fossil energy and water (de Vries and de Boer, 2010). The European livestock sector, for example, is responsible for more than 90% of the total emission of ammonia in the European union, and the dairy sector is one of the major contributors (Eurostat, 2015). Production of one kg of milk, furthermore, requires about 2.3–5.3 MJ of fossil energy (Upton et al., 2015).

At present, several environmental indicators are adopted to quantify and benchmark the environmental performance of dairy

* Corresponding author at: Animal Production Systems group, Wageningen University, P.O. Box 338, 6700 AH Wageningen, the Netherlands.

E-mail address: wenjuan.mu@wur.nl (W. Mu).

production systems, and to gain insight into potential improvement strategies. These environmental indicators are generally derived from a nutrient balance approach or a life cycle assessment (Oenema et al., 2003; Thomassen and de Boer, 2005; Yan et al., 2011).

A nutrient balance (NB) computes the difference in nutrients entering and leaving a system (Oenema et al., 2003), and allows computation of environmental indicators, such as nutrient use efficiency or nutrient loss per ha of land. An NB generally focusses on the nutrients nitrogen (N) and phosphorus (P), because these are the major nutrients that limit crop growth, and their losses can cause environmental problems (Oenema et al., 2003; Gourley et al., 2012). An NB of a dairy production system generally is computed at farm level. Indicators derived from an NB at farm level, however, do not include nutrient losses related to the production of farm inputs, such as purchased concentrates. Mu et al. (2016) demonstrated that an NB at chain level (i.e. cradle-to-farm gate) should be used to benchmark nutrient losses of dairy systems when differences in on-farm losses between systems are small, and pre-farm losses related to e.g. production of purchased concentrates, are relatively important.

Although environmental indicators derived from an NB appear to be useful to gain insight into the nutrient losses to the environment, generally they do not specify the type of losses, nor the environmental impact associated with those losses, such as the impact on acidification or climate change. Furthermore, a NB neglects certain environmental impact categories, such as the use of natural resources like fossil energy or land (Thomassen and de Boer, 2005).

Life cycle assessment (LCA) is an internationally accepted and standardized method (ISO 14040, ISO 14041, ISO 14042, ISO 14043) that quantifies the potential environmental impact related to emissions of pollutants to air, water and soil, and the use of resources during the entire life of a product. Thus, contrary to an NB approach, an LCA specifies the type of losses, as well as the associated environmental impact. Over the past few years the number of LCAs on dairy products have increased enormously (e.g. Cederberg, 1998; Thomassen and de Boer, 2005; Yan et al., 2011). However, studies suggested that collection of data required to perform an LCA appears difficult and is more time consuming than, for example, performing an NB (Thomassen and de Boer, 2005).

To benchmark the environmental performance of large groups of dairy production systems, there is a need for a set of sustainability indicators that does not require an excessive amount of data, and provides insights into the wider environmental impact of a system (Bélanger et al., 2012). Exploring correlations between various indicators can help to identify such a set of indicators (Lebacqz et al., 2013). Previous studies have mainly focused on correlations between indicators within LCA (Berger and Finkbeiner, 2011; Laurent et al., 2012; Röös et al., 2013). Berger and Finkbeiner (2011), for example, analysed correlation between several environmental indicators derived from an LCA on a hundred different materials (i.e. grouped into four categories 1) ore, metal, alloys; 2) monomers and polymers; 3) organic intermediates; 4) inorganic intermediates). They concluded that to compare the environmental performance of these materials, the number of indicators can be reduced because of strong correlations between several of the resource-oriented indicators. Laurent et al. (2012) analysed correlations between the carbon footprint and thirteen other LCA impact categories of about 4000 different products and concluded that solely relying on the carbon footprint as environmental indicator could result in overlooking other important environmental impacts. Röös et al. (2013), however demonstrated that the carbon footprint generally can act as an indicator for acidification and eutrophication potential of different types of meat (i.e., pork, chicken and beef). Results are explained by the importance of nitrogen losses, contributing to

both eutrophic and acidifying substances as well as to greenhouse gas emissions in the form of nitrous oxide.

So far, no study examined correlations between environmental indicators within dairy production, or included indicators derived from an NB. The purpose of our study, therefore, is to explore correlations between NB and LCA indicators, in order to identify an effective set of indicators that can be used as a proxy for the environmental performance of dairy systems. Such a set of indicators can be used, for example, to benchmark dairy farms.

2. Material and methods

2.1. Data

To identify an effective set of indicators to benchmark the environmental performance of dairy farms, we used farm data from Dairyman. Dairyman was a project in the INTERREGIVB program co-funded by the European Regional Development Fund that aimed to improve regional prosperity through better resource utilization on 113 dairy farms and stakeholder cooperation (Dairyman, 2010).

When exploring correlations between environmental indicators, we should avoid using data from contrasting production systems, because systematic differences in environmental impacts between systems bias the correlation analysis. We therefore selected 55 specialised dairy farms from Dairyman and determined their environmental performance using different indicators for 2010. We defined specialised farms as farms that have less than 5% non-dairy purpose animals, and less than 10% of their agricultural area in use for non-dairy purpose activities. The amount of energy, land and fertilizers used for non-dairy purposes was based on farmers' estimate and excluded from the data set. These 55 dairy farms are from different countries and regions (i.e. Netherlands, Ireland, Belgium (Flanders, Wallonia), France (Brittany), Germany (Baden, Württemberg) and Luxembourg) and differ in farm characteristics (Table 1).

2.2. System boundaries

Fig. 1 illustrates the system boundaries for the NB and LCA approach. For both approaches, system boundaries are from cradle-to-farm gate. For the NB, we included all on-farm activities as well as the production of purchased feed products, but production of other farm inputs were excluded because their contribution to nutrient losses was assumed to be negligible (Mu et al., 2016). In case of LCA, we considered on-farm processes, e.g. manure management, milk and feed production, and off-farm processes, e.g. fertilizer and feed production. Pesticides and water usage were not considered due to lack of data. Capital goods (buildings and machinery) were also excluded because their contribution to the environmental impact of dairy production is assumed to be low (Cederberg, 1998).

2.3. Nutrient balance

In this study, we used the method of Mu et al. (2016) to estimate N and P surpluses at chain level for each of the 55 specialised dairy farms. We first determined a NB at farm level, which equals the difference in nutrients entering and leaving the farm. Computation of a farm level NB requires data about the quantity and nutrient content of farm inputs and outputs, and data on stock changes of e.g. concentrates, roughages, animals and fertilizers on the farm. We calculated net inputs or net outputs of products that were both purchased and sold, such as animals. For the case of manure, we subtract manure outputs from the fertilizer input. The chain level NB was subsequently calculated by summing up the NB at farm

Table 1
General characteristics of the 55 dairy farms.

Country (#)	Farm area (ha)	Dairy cows (#)	Milk production (kg cow ⁻¹ year ⁻¹)	Concentrate (kg cow ⁻¹ year ⁻¹)
Belgium (15)	61(16.0)	82(23.8)	8053(1473.7)	1547(797.8)
France (2)	72(1.4)	77(21.7)	7192(794.7)	765.5(538.1)
Germany (4)	154(74.3)	102(17.4)	8866(1396.9)	2139(824.9)
Ireland (19)	77(28.3)	109(29.7)	5304(646.7)	841(180.6)
Luxembourg (1)	78(N.A.)	49(N.A.)	4826(N.A.)	710(N.A.)
Netherlands (14)	60(24.1)	126(35.1)	8705(752.5)	2066(245.0)

N.A. = not applicable.

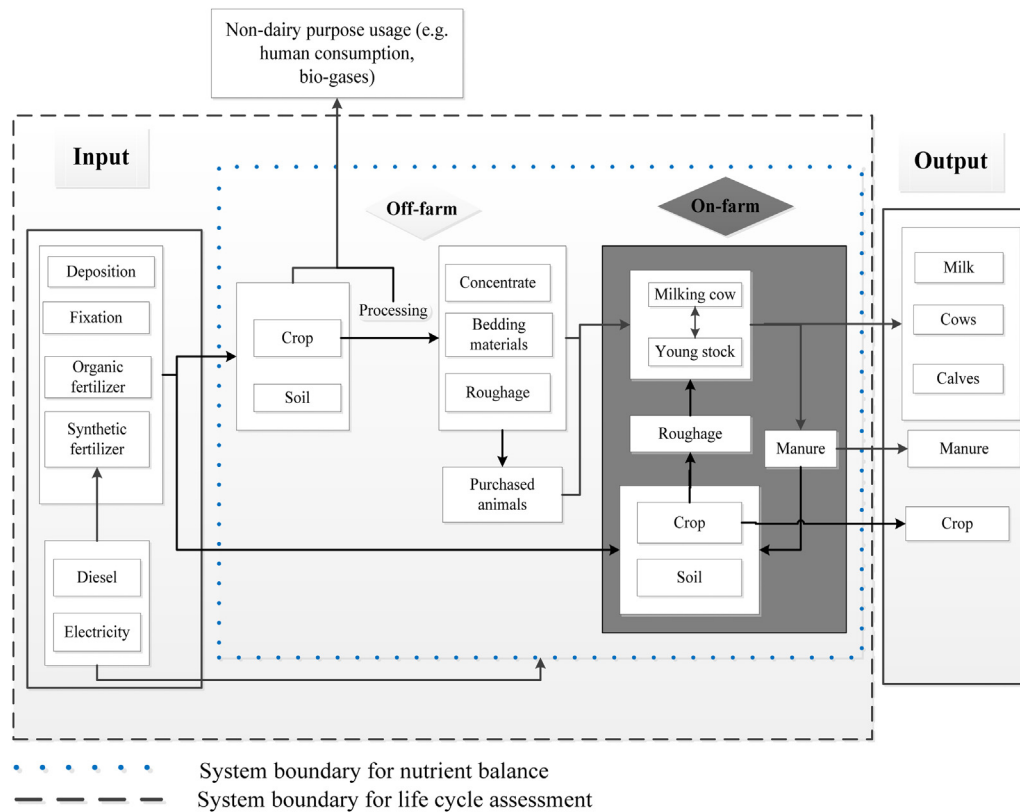


Fig. 1. System boundaries.

level and the NB related to production of purchased feed products. Unlike at farm level, nutrient surplus at chain level cannot be expressed per ha, because purchased feed ingredients are produced at different locations and different countries. Computing one average value for nutrient surplus per ha, therefore, would be misleading because it enables compensation of poor results for one of the ingredients by another, and does not provide insight into the local impact of any of the (crop) farms under study. The NB for N and P, therefore, were expressed per ton of fat-and-protein-corrected milk (FPCM), i.e. milk that is corrected to a fat content of 4.0% and a protein content of 3.3% using the formula: FPCM (ton) = Milk (ton) $\times$ $[0.337 + 0.116 \times \text{Fat (\%)} + 0.06 \times \text{Protein (\%)}]$ (Product Board Animal Feed, 2008)). Nutrient surpluses were allocated to milk and meat based on the relative economic value (i.e. economic allocation).

2.4. Life cycle assessment

This LCA study focussed on the use of scarce resources land and energy, and on the emission of pollutants methane (CH₄), nitrous oxide (N₂O), carbon dioxide (CO₂), ammonia (NH₃), mono-nitrogen oxides (NO_x), sulphur dioxide (SO₂), nitrate (NO₃⁻) and phosphate (PO₄³⁻). The impact categories considered were land use, fossil

energy use, climate change, acidification, freshwater eutrophication, and marine eutrophication (Table 2).

2.4.1. Inventory analysis

Land use related to on-farm processes was based on farm data (Dairyman, 2010), whereas energy use related to on-farm processes was based on farm data (Dairyman, 2010) and fixed factors from Dairyman (2010, based on Gac et al., 2010). The following fixed factors were used to assess energy use related to milking and on-farm crop production: electricity use during milking (65 kWh per 1000 l of milk), electricity use for crop production (37 kWh per ha), and diesel use for crop production (128 l per ha). In addition, diesel use for forage and pasture production was dependent on the area of maize silage (77 l per ha for areas with <5% maize silage; 94 l per ha for areas with 5–25% maize silage; and 128 l per ha for areas with >25% maize silage) (Dairyman, 2010; based on Gac et al., 2010).

On-farm CH₄ emissions relate to enteric fermentation and manure management. In this study, CH₄ emissions were based on IPCC (2006) Tier 2. Enteric CH₄ was assumed to equal 6.5% of the average gross energy (GE) intake of the dairy herd. GE intake was estimated based on farm data and dependent on energy requirements for growth, maintenance, activity, and milk production. Emissions of CH₄ from manure management were based on the

Table 2
Contributing elements and characterisation factors for selected impact categories.

Impact category	Unit	Contributing elements	Characterization factors	References
Land use	m ²	Land occupation		Guinee et al. (2002)
Energy use	MJ	Energy		
Climate change	Kg CO ₂ -eq	CO ₂	1	IPCC (2013)
		CH ₄ (biogenic)	28	
		CH ₄ (fossil)	30	
		N ₂ O	265	
Acidification	Kg SO ₂ -eq	SO ₂	1	ReCiPe (2008)
		NH ₃	2.45	
		NO _x	0.56	
Freshwater eutrophication	Kg P-eq	P	1	ReCiPe (2008)
		PO ₄ ³⁻	0.33	
Marine eutrophication	Kg N-eq	N	1	ReCiPe (2008)
		NH ₃	0.092	
		NO _x	0.039	
		NO ₃ ⁻	0.23	

volatile solid excretion rate of the dairy herd, and type of manure system. Volatile solid excretion was estimated based on the GE intake and an assumed feed digestibility of 75% for dairy cows, and 65% for young stock (IPCC, 2006). CH₄ per kg volatile solid excreta was calculated by multiplying the maximum amount of 0.36 kg CH₄/kg (dairy cow) or 0.12 kg CH₄/kg (young stock) with a system specific conversion factor (i.e., 19% for manure in pit storage, 2% for solid manure, 1% for manure excreted on pasture).

On-farm N₂O emissions occur via direct and indirect pathways. Direct N₂O emissions result from manure management and N application to the field, whereas indirect N₂O emissions results from volatilization of NH₃, NO_x, and leaching of NO₃⁻ (IPCC, 2006). Direct N₂O emissions from manure storage were based on the N excretion multiplied by a management specific emission factor (i.e., 0.002 kg N₂O-N per kg N for manure in pit storage, and 0.005 kg N₂O-N per kg N for solid manure; IPCC, 2006). N excretion was based on N intake (i.e., calculated from the dry matter intake and N content of the diet) and the amount of N retained by the dairy herd for milk production and growth. Direct N₂O emissions from N application via manure, synthetic fertilizer, and crop residues, and all indirect N₂O emissions were based on IPCC Tier 1 (IPCC, 2006).

On-farm NH₃ emissions relate to manure management and N application to the field, and were based on the Tier 2 approach of the European Environmental Agency (EEA) air pollutant emission inventory guidebook (EEA, 2013). Emissions from manure management were estimated based on the total ammonia nitrogen (TAN) content of manure, which was assumed to be 60% of the total N excretion. Emission factors in kg NH₃-N per kg TAN were 0.2 for slurry and 0.19 for solid manure in stables; 0.20 for slurry and 0.27 for solid manure in storage; 0.55 for slurry and 0.79 for solid manure after application; and 0.10 (dairy cows) and 0.06 (young stock) for manure excreted during grazing. Emissions of NH₃ related to the application of synthetic fertilizer were based on fertilizer specific emission factors (EEA, 2013).

On-farm emissions of NO_x and SO₂ related to the combustion of diesel were based on Eco-invent (2010).

Due to lack of detailed soil information, we estimated leaching of NO₃⁻ and PO₄³⁻ based on the N and P balance of the farm. This approach is often applied in environmental impact assessment studies (van Middelaar et al., 2013), but could lead to an overestimation of leaching because, potential nutrient build-up in soils is neglected. Leaching of NO₃⁻ was calculated by subtracting NH₃-N, N₂O-N and NO_x-N from the total N surplus of the farm, whereas for PO₄³⁻, the total P surplus was assumed to leach off. Phosphate fixation in the soil was assumed to be negligible in this study because we do not have detailed soil information. Uwizeye et al. (2016), however, showed that P saturation can vary a lot across EU

soils. We, therefore, may have overestimated phosphate leaching on some of the farms.

Off-farm processes contributing to use of resources and emissions of pollutants are the production and transportation of purchased feed (e.g. concentrates, roughages, by-products), the production and transportation of fertilizers and pesticides, and the production and transportation of energy sources. Impacts related to production and transportation of purchased feed were based on FeedPrint (2015), whereas impacts related to production of synthetic fertilizers and energy sources were based on Eco-invent (2010). Production and transportation of pesticides, and transportation of fertilizers and energy sources were not included. The contribution of these processes was assumed to be minor.

2.4.2. Impact assessment

Resources and emissions were classified by the following impact categories: land use, fossil energy use, climate change, acidification, freshwater eutrophication and marine eutrophication. Climate change relates to the emission of greenhouse gases, including CO₂, CH₄ and N₂O. Acidification of terrestrial and aquatic ecosystems relates to the emission of acidifying gases, including SO₂, NH₃ and NO_x (i.e., NO and NO₂). Eutrophication of terrestrial and aquatic ecosystems relates to leaching of eutrophic substrates, including NO₃⁻ and PO₄³⁻, and to the emission of eutrophic gases, including NH₃ and NO_x. We distinguish between freshwater and marine eutrophication, as suggested by ReCiPe (2008), and assumed that N is the limiting nutrient in costal water, and P in marine water.

The characterisation factors for the different pollutants per impact categories are listed in Table 2. The functional unit is one ton of FPCM. Economic allocation was used to allocate the environmental impact between the different outputs in case of a multifunctional process, i.e. impacts were allocated based on the relative economic value of the outputs.

2.5. Statistical analysis

First, Pearson and Spearman correlations were calculated between eight indicators at the chain level, i.e. N surplus, P surplus, land use, fossil energy use, global warming potential (GWP), acidification potential (AP), freshwater eutrophication potential (FEP) and marine eutrophication potential (MEP), all expressed per kg FPCM. These correlations offer a first impression of the strength of relationships for pairs of indicators. Spearman rank correlations were added as a simple check upon the presence of outlying, influential data.

Second, we applied multiple regression analyses to explore the relationships between each indicator and remaining other indica-

tors. In this second step, indicators that strongly relate to other indicators will be omitted from consideration and a sub-set of indicators will remain that can be used as proxies. In separate multiple regression analyses, in turn each of the eight indicator was regressed on the other seven indicators. The proportion of variance, the so-called coefficient of determination R^2 , which equals the square of the multiple correlation coefficient, was stored. Indicators were ranked based on their associated R^2 values. If the indicator with the highest rank had an R^2 larger than 0.64, this indicator was eliminated, because it can be predicted well by the other indicators. Threshold value 0.64 was chosen because it corresponds to a multiple correlation of 0.8, which is considered to be high (Cohen, 1988). Subsequently, the procedure was repeated for the remaining seven indicators, regressing each indicator in turn upon the other six indicators. Again, the indicator, now among the seven remaining indicators, with the highest R^2 value was eliminated when the associated R^2 exceeded 0.64. This procedure was repeated until none of the remaining indicators was associated with an R^2 above 0.64. The remaining indicators are the set of indicators to be used as proxies.

Finally, we investigated canonical correlation between the set of all eliminated indicators and the set of remaining indicators to test if the set of remaining indicators can be used as a proxy for the eliminated indicators. This selection procedure does not take into account any practical considerations, e.g. some indicators being more easily observed than others. Referring to the set of selected indicators as the optimal set, an alternative set of indicators associated with multiple and canonical correlations similar to those of the optimal set can be considered for use as proxies as well, when some of the indicators in this alternative set are more appealing from a practical point of view. In this respect the optimal set is considered to be a standard for comparison with other alternative sub-sets. All calculations were performed with SPSS (SPSS, 2011).

3. Results

3.1. Nutrient balance

Indicators derived from the NB are presented in Table 3. On average, nutrient surplus at chain level are 14.4 g N/kg FPCM, and 1.3 g P/kg FPCM. An important part of the nutrient surplus relate to on-farm processes; 84% in case of N and 41% in case of P. Variation in nutrient surplus among farms is large and mainly relate to variation in on-farm nutrient surplus. Results are in line with results found in literature (Gourley et al., 2012; Buckley et al., 2013; Leip et al., 2013; Mu et al., 2016).

3.2. Life cycle assessment

Indicators derived from the LCA are presented in Table 4. Producing one kg FPCM requires on average 1.3 m² land and 2.8 MJ

Table 3
Indicators derived from the nutrient balance.

Indicator	Unit	Mean	SD	Minimum	Maximum
<i>N surplus</i>					
On-farm	g N × kg FPCM ⁻¹	12.2	5.3	3.5	25.0
Off-farm	g N × kg FPCM ⁻¹	2.3	1.3	0.4	6.2
Total	g N × kg FPCM ⁻¹	14.4	5.2	4.9	27.0
<i>P surplus</i>					
On-farm	g P × kg FPCM ⁻¹	0.5	0.7	0.0	3.0
Off-farm	g P × kg FPCM ⁻¹	0.8	0.4	0.1	1.9
Total	g P × kg FPCM ⁻¹	1.3	0.9	0.1	4.9

Table 4
Indicators derived from the LCA.

Indicator	Unit	Mean	SD	Minimum	Maximum
<i>Land use</i>					
On-farm	m ² × kg FPCM ⁻¹	0.8	0.4	0.3	2.3
Off-farm		0.4	0.2	0.1	1.3
Total		1.3	0.4	0.7	2.8
<i>Energy use</i>					
On-farm	MJ × kg FPCM ⁻¹	0.5	0.1	0.2	1.0
Off-farm		2.4	0.8	1.4	4.7
Total		2.8	0.8	1.8	5.1
<i>GWP^a</i>					
On-farm	kg CO ₂ eq × kg FPCM ⁻¹	1.0	0.2	0.7	1.5
Off-farm		0.2	0.1	0.0	0.5
Total		1.2	0.2	0.8	1.8
<i>AP^a</i>					
On-farm	g SO ₂ eq × kg FPCM ⁻¹	22.5	6.1	7.2	38.2
Off-farm		3.6	1.4	1.1	8.5
Total		26.1	6.2	10.3	42.4
<i>FEP^a</i>					
On-farm	g P eq × kg FPCM ⁻¹	0.6	0.7	0.0	3.1
Off-farm		0.6	0.3	0.1	1.9
Total		1.1	0.9	0.1	5.0
<i>MEP^a</i>					
On-farm	g N eq × kg FPCM ⁻¹	5.3	3.8	0.7	16.1
Off-farm		2.8	1.2	0.7	6.9
Total		8.1	4.0	1.7	19.4

^a GWP (global warming potential), AP (acidification potential), FEP (freshwater eutrophication potential), MEP (marine eutrophication potential).

energy, and results in a GWP of 1.2 kg CO₂-eq, an AP of 26.1 g SO₂-eq, a FEP of 1.1 g P-eq and a MEP of 8.1 g N-eq. On-farm activities have a larger impact on land use, GWP, AP and MEP compared to off-farm activities. Off-farm activities, however, use more energy than on-farm activities. The results of our study are in line with results found in literature (de Boer, 2003; Cederberg and Flysjö, 2004; O'Brien et al., 2012; Upton et al., 2013).

Table 5
Pearson correlation coefficient matrix of the environmental indicators.^{***}

	N surplus	P surplus	Land use	Energy use	GWP ^a	AP ^a	FEP ^a	MEP ^a
N surplus	–	0.30*	0.24	0.59**	0.68**	0.64**	0.36**	0.92**
P surplus		–	0.24	0.21	0.15	0.45	0.93**	0.31*
Land use			–	0.69**	0.61**	0.54**	0.41**	0.20
Energy use				–	0.78**	0.72**	0.37**	0.50**
GWP ^a					–	0.81**	0.28*	0.51**
AP ^a						–	0.19	0.42**
FEP ^a							–	0.41**
MEP ^a								–

* Correlation is significant at the 0.01 level (2-tailed).

** Correlation is significant at the 0.05 level (2-tailed).

*** The results of the Pearson correlation are in line with the results of Spearman Rho's correlation, we therefore only presented the results of the Pearson correlation coefficients here.

^a GWP (global warming potential), AP (acidification potential), FEP (freshwater eutrophication potential), MEP (marine eutrophication potential).

Table 6

Multiple correlations between the set of indicators to be used as a proxy to benchmark environmental performance i.e., P surplus, land use, acidification potential, marine eutrophication potential and each of the eliminated indicators, and the canonical correlation between both sets.

Eliminated indicators	Correlation coefficient [*]
Multiple correlations	
N surplus	0.96
FEP ^a	0.95
GWP ^a	0.84
Energy use	0.82
Canonical correlation	
N surplus, FEP, ^a GWP, ^a energy use	0.98

^{*} All correlation coefficients are significant at the 0.05 level (2-tailed).

^a FEP (freshwater eutrophication potential), GWP (global warming potential).

Table 7

Multiple correlations between the new set of indicators to be used as a proxy to benchmark environmental performance (i.e. N surplus, P surplus, land use, energy use) and each of the eliminated indicators, and the canonical correlation between both sets.

Eliminated indicator	Correlation coefficient [*]
Multiple correlation	
FEP ^a	0.95
MEP ^a	0.91
GWP ^a	0.83
AP ^a	0.79
Canonical correlation	
FEP, MEP, GWP, AP	0.97

^{*} All correlation coefficients are significant at the 0.05 level (2-tailed).

^a FEP (freshwater eutrophication potential), MEP (marine eutrophication potential), GWP (global warming potential), AP (acidification potential).

3.3. Statistical analysis

Correlations are presented in Table 5. A strong correlation was found between N surplus and MEP ($r=0.92$), and between P surplus and FEP ($r=0.93$). Furthermore, strong or marked correlation were found between GWP and AP ($r=0.81$), GWP and N surplus ($r=0.68$), between GWP and energy use ($r=0.78$), and GWP and land use ($r=0.61$), between energy use and land use ($r=0.69$) and energy use and AP ($r=0.72$), and between AP and N surplus ($r=0.64$).

Through repeated use of multiple regression analysis a set of indicators was selected i.e. MEP, P surplus, land use, and AP that can be used as a proxy to benchmark environmental performance. The multiple correlations between the set of indicators to be used as a proxy and each of the eliminated indicators are shown in Table 6. The set of indicators to be used as a proxy has a strong correlation with N surplus ($r=0.96$), FEP ($r=0.95$), GWP ($r=0.84$), and energy use ($r=0.82$). The canonical correlation between the set of indicators selected as proxies and the set of eliminated indicators is also high ($r=0.98$; Table 6).

The set of indicators selected to be used as a proxy to benchmark the environmental performance of dairy farms contains four indicators of which two (MEP and AP) require more data and complex calculation methods than the other two (P surplus and land use). Since MEP is highly correlated with N surplus, and energy use is also correlated with AP, we further explored the possibility to substitute MEP with N surplus, and substitute AP with energy use, which can reduce data requirements and simplify computations. The multiple correlations between this new set of indicators, i.e. N surplus, P surplus, land use, energy use, and each of the eliminated indicators, i.e. MEP, FEP, GWP, AP, and the canonical correlation between the two sets are presented in Table 7. The alternative set of indicators to be used as a proxy are also strongly correlated with the

eliminated indicators FEP ($r=0.95$), MEP ($r=0.91$), GWP ($r=0.83$) and AP ($r=0.79$). The canonical correlation between the two sets is high as well ($r=0.97$). Therefore, the alternative set of indicators does not differ much from the set of indicators that was initially selected (Table 6). We, therefore, recommend that N surplus, P surplus, land use and energy use will be used as proxies to benchmark the environmental performance of dairy farms.

4. Discussion

In recent years, the key question in the field of sustainability assessment of livestock production systems has shifted from “how do we develop an indicator?” to “which indicators can we use?”, highlighting the importance of indicator selection (Bockstaller et al., 2008; Lebacqz et al., 2013). Selecting an effective set of indicators will depend on factors, such as the intended purpose and the target group. Complex indicators that require extensive information and expert knowledge to provide sound environmental performance evaluations are generally not suitable for benchmarking purposes (Bélanger et al., 2012). In the case of benchmarking, indicators should be simple, measurable, accessible, relevant and timely (Lockie et al., 2002). To be more specific, an effective set of indicators to benchmark dairy systems are defined as indicators that are able to assess the major environmental performance of the farms, and able to provide early warning of potential environmental problems; meanwhile, this set of indicators needs to be monitored and collected in a time and cost efficient way, and to be interpreted easily by decision makers. Benchmarking results should be comprehensible and easily applicable by decision makers, such as farmers and policy makers; they should provide insight into the environmental hotspots of the production systems and guide strategic (e.g. change in production system) or operational (e.g. change in feeding strategy) decisions to reduce negative environmental impacts of dairy farming. However, it should be noted that by using a selected set of indicators, there are risks that environmental impacts that are not represented by the selected set are overlooked in the benchmarking process.

In this study, N surplus, P surplus, land use and energy use were identified as the effective set of indicators to benchmark environmental performance of dairy farms and represent also MEP, FEP, AP and GWP. MEP logically can be represented by N surplus, because leaching of NO_3^- is the major contributing element to MEP and is highly correlated with N surplus. The same holds for FEP, which is caused by leaching of PO_4^{3-} and, therefore, can be predicted by P surplus. Meanwhile, AP can be represented by N surplus and energy use because N surplus include emissions of NH_3 and NO_x , and production and combustion of energy sources contribute to the emission of SO_2 , NO_x , which are all elements contributing to AP. GWP can be represented by N surplus and energy use because N surplus include the emission of N_2O , and production and combustion of energy sources contribute to the emission of CO_2 . The GWP, however, is determined not only by the emissions of N_2O and CO_2 , but also by the emission of CH_4 , mainly related to enteric fermentation, and contributing about 50% to the total GWP of milk.

Due to the simplicity of the method for quantifying some of the emissions, variation in environmental impacts between farms was limited for some indicators. Variation in CH_4 per unit of milk, for example, was determined mainly by milk production levels (as were the other indicators) and not so much by differences in dietary composition of the herd or other feed characteristics. As a result, correlations between GWP and N surplus and energy use were quite high, even though enteric CH_4 is not related to N surplus and energy use. If information on, e.g., diet composition would be available, we could assess CH_4 emissions with a more detailed approach (e.g. Tier 3), which could have affected the selection of the final set of indica-

tors. The methods that were used in this study, however, are widely accepted and generally used in current research and practices.

This study used data from specialised dairy farms to identify a set of indicators to benchmark the environmental performance of farms. We deliberately selected specialized farms only to avoid bias in our correlation analysis resulting from systematic differences in environmental impacts between contrasting systems (i.e. mixed crop-livestock systems versus specialized systems). Plotting the environmental impacts demonstrated the absence of systematic differences between countries or systems. Based on current methodologies, we expect that the reduced set of indicators can also be used to benchmark environmental performance of other types of dairy systems, because like we explained previously, the mechanism behind the relationships of the indicators is irrelevant of the type of dairy system; N surplus and P surplus can be used as good proxies for estimating e.g. leaching of NO_3^- and PO_4^{3-} that contribute to MEP and FEP, and the combination of N surplus and energy use can be used as a proxy for estimating NH_3 , NO_x , SO_2 , N_2O and CO_2 , which contribute to AP and GWP. However, with more advanced methodology, e.g. to estimate emissions of CH_4 related to enteric fermentation, we might end up with a different set of indicators. Future research, however, would be required to verify this hypothesis.

Unlike some indicators in LCA that are difficult to quantify, our set of indicators are technically and financially feasible to be collected on a regular basis because of the transparency in the usage of the input parameters. Indicators are also relatively easy to communicate with farmers and policy makers, which facilitates identification of improvement options. For example, if the N surplus on a specific farm is relatively high, the farmer can easily oversee which N inputs contribute most (Thomassen and de Boer, 2005). The reduced set of indicators can represent the full set of indicators included in this study. One indicator is not regarded to be more important than the others because they all represent different environmental impacts. The importance of the indicators may differ between farms or regions, and could depend on the decision of stakeholders or policy makers. The indicators we examined in this study do not provide detailed information on the local impacts of the emissions. Eutrophication potential, for example, was based on generic impact factors because local specific impact factors are not available at the global level. In addition, it should be noted that the selected set of indicators only represents the limited set of indicators that was included in this study. Impact on biodiversity and water use, for example, were not assessed. Further research is required to assess correlations between a larger set of indicators.

5. Conclusion

The objective of this study was to identify an effective set of indicators that can be used to benchmark the environmental performance of dairy systems. N surplus, P surplus, energy use and land use were selected to represent a broader set of indicators also including global warming potential, acidification potential, freshwater eutrophication potential and marine eutrophication potential. This effective set of indicators can be used to benchmark farms on some major environmental impacts, and they can be monitored and collected in a time and cost-effective way, and can be interpreted easily by decision makers. Other important environmental impacts, such as biodiversity and water use, however, should not be overlooked.

Acknowledgements

This research is part of the Wageningen UR strategic programme 'Mapping for sustainable intensification', 2012–2016, funded by

the Strategic Funds of Wageningen UR, and the WASS and WIAS Graduate Schools of Wageningen University.

References

- Bélanger, V., Vanasse, A., Parent, D., Allard, G., Pellerin, D., 2012. Development of agri environmental indicators to assess dairy farm sustainability in Quebec Eastern Canada. *Ecol. Indic.* 23, 421–430.
- Berger, M., Finkbeiner, M., 2011. Correlation analysis of life cycle impact assessment: indicators, measuring resource use. *Int. J. Life Cycle Assess.* 16, 74–81.
- Bockstaller, C., Guichard, L., Makowski, D., Aveline, A., Girardin, P., Plantureux, S., 2008. Agri-environmental indicators to assess cropping and farming systems. A review. *Agron. Sustain. Dev.* 28, 139–149.
- Buckley, C., Murphy, P., Wall, D., 2013. Farm-Gate N and P Balances and Use Efficiencies across Specialist Dairy Farms in the Republic Ireland. Working paper 13-WP-RE-02 Teagasc, Rural Economy and Development Programme.
- Cederberg, C., Flysjö, A., 2004. Life Cycle Inventory of 23 Dairy Farms in South-Western Sweden. SIK-report 728. Swedish Institute for Food and Biotechnology, SIK, Gothenburg, Sweden.
- Cederberg, C., 1998. Life Cycle Assessment of Milk Production – a Comparison of Conventional and Organic Farming SIK. Report No. 643. The Swedish Institute for Food and Biotechnology, Göteborg, Sweden.
- Cohen, J., 1988. Statistical Power Analysis for the Behavioral Sciences, 2nd ed. Lawrence Erlbaum, Associates Inc, Abingdon.
- Dairyman, 2010. <<http://www.interregdairyman.eu/en/dairyman.htm>>.
- European Environment Agency (EEA), 2013. EMEP/EEA Air Pollutant Emission Inventory Guidebook, <http://www.eea.europa.eu/publications/emep-eea-guidebook-2013> (accessed 08.03.16.).
- Eco-invent, 2010. Eco-Invent Data v2.2. Final Reports Eco-Invent. Swiss Centre for Life Cycle Inventories, Dübendorf.
- Eurostat, 2015. http://ec.europa.eu/eurostat/statistics-explained/index.php/Air_pollution_statistics (accessed 08.03.16.).
- FAO, 2006. FAO World Agriculture: Towards 2030/2050 Interim Report. Food and Agriculture Organisation, Rome.
- FAOSTAT, 2013. FAO statistical databases: agriculture, fisheries, forestry, nutrition crop statistics. FAO—Food and Agriculture Organization of the United Nations, Rome.
- FeedPrint, 2015. <http://www.wageningenur.nl/en/show/FeedPrint-Calculate-CO2-per-kilogram-meat-milk-or-eggs.htm> (accessed 03.08.16.).
- Gac, A., Deltour, L., Cariolle, M., Dolle, J.B., Espagnol, S., Flenet, F., Guinand, N., Lagadec, S., Le Gall, A., Lellahi, A., Malaval, C., Ponchant, P., Tailleur, A., 2010. Guide méthodologique pour l'estimation des impacts des activités agricoles sur l'effet de serre. Réalisé dans le cadre du projet «Gaz à Effet de Serre et Stockage de Carbone en exploitations agricoles» (CASDAR n(6147).
- Gerber, P.J., Steinfeld, H., Henderson, B., Mottet, A., Opio, C., Dijkman, J., Faluccci, A., Tempio, G., 2013. Tackling Climate Change Through Livestock – a Global Assessment of Emissions and Mitigation Opportunities. Food and Agriculture Organization of the United Nations, Rome.
- Gourley, C.J.P., Dougherty, W.J., Weaver, D.M., Aarons, S.R., Awty, I.M., Gibson, D.M., Hannah, M.C., Smith, A.P., Peverill, K.L., 2012. Farm-scale nitrogen, phosphorus, potassium and sulphur balances and use efficiencies on Australian dairy farms. *Anim. Prod. Sci.* 52, 929–944.
- Guinee, J.B., Gorree, M., Heijungs, R., Huppes, G., Kleijn, R., de Koning, A., Van Duin, R., Huijbregts, M.A.J., 2002. Handbook on life cycle assessment. In: Operational Guide to the ISO Standards. Kluwer Academic Publishers, Dordrecht.
- IPCC, 2006. Intergovernmental Panel on Climate Change. Guidelines for national greenhouse gas inventories. In: Eggleston, H.S., Buendia, L., Miwa, K., Ngara, T., Tanabe, K. (Eds.), Agriculture, Forestry and Other Land Use. Prepared by the National Greenhouse Gas Inventories Program, vol. 4. IGES, Japan.
- IPCC, 2013. Climate Change 2013: the Physical Science Basis. Working Group I Contribution to the Fifth Assessment Report of the Intergovernmental Panel on Climate Change. Chapter 8: Anthropogenic and Natural Radiative Forcing.
- ISO 14040, 1997. Environmental Management – Life Cycle Assessment: Principles and Framework. European Committee for Standardization (CEN), Brussels, Belgium.
- ISO 14041, 1998. Environmental Management – Life Cycle Assessment: Goal and Scope Definition and Inventory Analysis. European Committee for Standardization (CEN), Brussels, Belgium.
- ISO 14042, 2000. Environmental Management – Life Cycle Assessment: Life Cycle Impact Assessment. European Committee for Standardization (CEN), Brussels, Belgium.
- ISO 14043, 2000. Environmental Management – Life Cycle Assessment: Life Cycle Interpretation. European Committee for Standardization (CEN), Brussels, Belgium.
- Laurent, A., Olsen, S.I., Hauschild, M.Z., 2012. Limitations of carbon footprint as indicator of environmental sustainability. *Environ. Sci. Technol.* 46, 4100–4108.
- Lebacqz, T., Baret, P.V., Stilmant, D., 2013. Sustainability indicators for livestock farming. A review. *Agron. Sustain. Dev.* 33, 311–327.
- Leip, A., Weiss, F., Lesschen, J.P., Westhoek, H.J., 2013. The nitrogen footprint of food products in the European Union. *J. Agr. Sci.*, 1–44.
- Lockie, S., Lawrence, G., Dale, A., Taylor, B., 2002. Capacity for change: testing a model for the inclusion of social indicators in Australia's land and water resources audit. *J. Environ. Plann. Manag.* 45, 813–826.

- Mu, W., van Middelaar, C.E., Bloemhof, J.M., Oenema, J., de Boer, I.J.M., 2016. Nutrient balance at chain level: a valuable approach to benchmark nutrient losses of milk production systems. *J. Clean. Prod.* 112, 2419–2428.
- O'Brien, D., Shalloo, L., Patton, J., Buckley, F., Grainger, C., Wallace, M., 2012. A life cycle assessment of seasonal grass-based and confinement dairy farms. *Agric. Syst.* 107, 33–46.
- Oenema, O., Kros, H., de Vries, W., 2003. Approaches and uncertainties in nutrient budgets: implications for nutrient management and environmental policies. *Eur. J. Agron.* 20, 3–16.
- Product Board Animal Feed, 2008. Tabellenboek Veevoeding 2008. Productschap Diervoeder, CVB, Den Haag, the Netherlands.
- Röös, E., Sundberg, C., Tidåker, P., Strid, I., Hansson, P.A., 2013. Can carbon footprint serve as an indicator of the environmental impact of meat production? *Ecol. Indic.* 24, 573–581.
- ReCiPe, 2008 A life cycle impact assessment method which comprises harmonised category indicators at the midpoint and the endpoint level, 1st ed. Report I: Characterisation, 2009. <http://www.lcia-recipe.net/> (accessed 08.03.16.).
- SPSS, 2011. Statistical Package for Social Sciences. SPSS Software Package 20.0 for Microsoft. Windows, Chicago.
- Thomassen, M.A., de Boer, I.J.M., 2005. Evaluation of indicators to assess the environmental impact of dairy production systems. *Agric. Ecosyst. Environ.* 111, 185–199.
- Upton, J., Humphreys, J., Groot Koerkamp, P.W.G., French, P., Dillon, P., De Boer, I.J.M., 2013. Energy demand on dairy farms in Ireland. *J. Dairy Sci.* 96, 6489–6498.
- Upton, J., Murphy, M., De Boer, I., Groot Koerkamp, P., Berentsen, P., Shalloo, L., 2015. Investment appraisal of technology innovations on dairy farm electricity consumption. *J. Dairy Sci.* 98, 898–909.
- Uwizeye, A., Gerber, A., Schulte, R., de Boer, I.J.M., 2016. A comprehensive framework to assess the sustainability of nutrient use in global livestock supply chains. Working paper.
- Yan, M.J., Humphreys, J., Holden, M.N., 2011. An evaluation of life cycle assessment of European milk production. *J. Environ. Manage.* 92, 372–379.
- de Boer, I.J.M., 2003. Environmental impact assessment of conventional and organic milk production. *Livest. Prod. Sci.* 80, 69–77.
- de Vries, M., de Boer, I.J.M., 2010. Comparing environmental impacts for livestock products: a review of life cycle assessments. *Livest. Sci.* 128, 1–11.
- van Middelaar, C., Cederberg, C., Vellinga, T., van der Werf, H.G., de Boer, I.J.M., 2013. Exploring variability in methods and data sensitivity in carbon footprints of feed ingredients. *Int. J. LCA* 18, 768–782.