Benchmarking the environmental performance of dairy farming systems

Wenjuan Mu

Thesis committee

Promotors

Prof. Dr I.J.M. de Boer
Professor of Animal Production Systems
Wageningen University & Research

Prof. Dr J.M. Bloemhof-Ruwaard
Professor of Operations Research and Logistics
Wageningen University & Research

Co-promotor

Dr C.E. van Middelaar
Post-doc researcher, Animal Production Systems Group
Wageningen University & Research

Other members

Prof. Dr M.K. van Ittersum, Wageningen University & Research
Dr J.W. Reijs, LEI Wageningen University & Research
Dr G. Emvalomatis, University of Dundee, Dundee, United Kingdom
Dr J. Upton, Irish Agriculture and Food Development Authority (Teagasc), Cork, Ireland

Benchmarking the environmental performance of dairy farming systems

Wenjuan Mu

Thesis

submitted in fulfilment of the requirements for the degree of doctor

at Wageningen University

by the authority of the Rector Magnificus,

Prof. Dr A.P.J. Mol,

In the presence of the

Thesis Committee appointed by the Academic Board

to be defended in public

on Wednesday 22 November 2017

at 1.30 p.m. in the Aula.

Contents

Abstract

Milk production has a major impact on the environment and competes increasingly for scarce resources. As the demand for milk is expected to increase, these issues are likely to worsen. Benchmarking the environmental performance of dairy farming systems offers the opportunity to identify best farm practices and to provide guidance for reducing the environmental impact. Currently, benchmarking is hampered by the lack of an effective method that results in a set of indicators that is easily quantifiable and detects variations in environmental performance between farms. The aim of this thesis, therefore, was to develop a sound method to benchmark the environmental performance of dairy farming systems. This thesis focuses is on specialized dairy farming systems in Europe.

The first challenge in benchmarking the environmental performance of dairy farming systems is to select a set of indicators that are relevant, measurable, valid, timely and understandable. Environmental indicators can be derived from various approaches, including a nutrient balance (NB) approach and a life cycle assessment (LCA). An NB is generally applied at farm level, and yields indicators that are relatively easy to quantify and communicate. We found that an NB at farm level can be used to benchmark dairy farming systems, if differences in on-farm losses are large and off-farm losses are relatively unimportant. Only if farms differ largely in the amount and/or type of purchased inputs, such as feed, the farm-based NB should be extended to a chain based NB or an LCA. An LCA, however, requires extensive data information, which can be difficult to collect. We, therefore, explored correlations between eight commonly used NB and LCA indicators with the system boundary from cradle-to-farm gate. We found that a set indicators, consisting of the nitrogen surplus, the phosphorus surplus, land use and energy use can be used as a proxy to benchmark the environmental performance of dairy farming systems, representing also global warming potential, acidification potential, freshwater eutrophication potential and marine eutrophication potential.

The second challenge in benchmarking the environmental performance of dairy farming systems is to cope with data uncertainties. We therefore first evaluated the effect of epistemic uncertainty on benchmarking the nitrogen use efficiency of dairy systems. We found that ranking of farms based on this single indicator is not possible when the epistemic uncertainty of parameters is large and differences in N use efficiency are small. We, furthermore, identified the most influential parameters (e.g. input of concentrates, mineral fertilizer) and found that reducing epistemic uncertainty of those parameters improved benchmarking results significantly. Afterwards, we demonstrated how to use fuzzy data envelopment analysis (DEA) to account for uncertainties of multiple indicators in benchmarking the eco-efficiency of dairy farming systems. With fuzzy DEA, the number of farms receiving the highest efficiency score was lower compared to standard DEA. In addition, fuzzy DEA identified a different set of peers than standard DEA. By taking uncertainty into account during the quantification processes, fuzzy DEA can contribute to increasing the reliability of results and prevent biased conclusions.

Exploring correlations between environmental indicators can facilitate decision-makers to derive an effective set of indicators that can be used as proxies for benchmarking. In addition, decision-makers should acknowledge the effect of epistemic uncertainty on benchmarking results. When setting up reference values for penalty, for example, this value should be based on a range rather than a single value in order to account for epistemic uncertainty.

Chapter 1

General introduction

1.1 Background

Milk and dairy products are consumed by billions of people around the world every day, because of their highly nutritious value to meet the need of the human body for especially protein and micronutrients, such as calcium, magnesium, and vitamin B_{12} (FAO, 2013). Milk production, furthermore, generates farm income, contributes to the gross domestic product of many countries, and contributes to food security and livelihoods of smallholders in countries like India, Pakistan and China (OECD,2016).

Milk production, however, also causes pressure on the environment, by its emissions to air, water, and soil, that contribute to, for example, climate change, eutrophication, acidification, and loss of biodiversity (Thomassen and De Boer, 2005). Climate change, induced by the emission of greenhouse gases (GHG), is one of the main challenges of today. The global dairy sector, producing both milk and meat, is responsible for about 30% of the anthropogenic GHG emissions related to livestock production (Gerber et al., 2013). The three main GHGs from milk production are carbon dioxide (CO_2), methane (CH_4), and nitrous oxide (N_2O). CO_2 is produced mainly by combustion of fossil energy during, e.g. transport or milking, and by burning and microbial decay of biomass related to changes in land use or crop management. CH_4 is produced mainly during enteric fermentation and anaerobic decomposition of organic material in manure; and N_2O is produced mainly through nitrification and denitrification during manure management and nitrogen fertilizer application. Besides the emission of GHGs, milk production results in the emission of other gases (e.g. ammonia) and substrates (e.g. nitrate or phosphate) leading to environmental impacts, such as acidification and eutrophication. Ammonia (NH_3) accounts for about 85% of the acidity equivalent emissions for livestock in Europe. The European livestock sector is responsible for more than 80% of the total agricultural emission of NH_3, and the dairy sector is one of the major contributors (Leip et al. 2015). In addition to NH_3, leaching of nitrate (NO_3^-) and phosphate (PO_4^{3-}) can deteriorate the quality of freshwater and coastal water, which can lead to the enhancement of water eutrophication (Leip et al., 2015). Furthermore, milk production threats biodiversity because of its impact on climate, air

and water pollution, land degradation and deforestation (Steinfeld et al., 2006), and it competes increasingly for scarce resources, such as land, water and fossil energy (De Vries and De Boer, 2010). As the population grows and becomes wealthier, milk production is expected to further increase (see Figure 1.1). Aforementioned environmental consequences related to milk production, therefore, are likely to amplify.
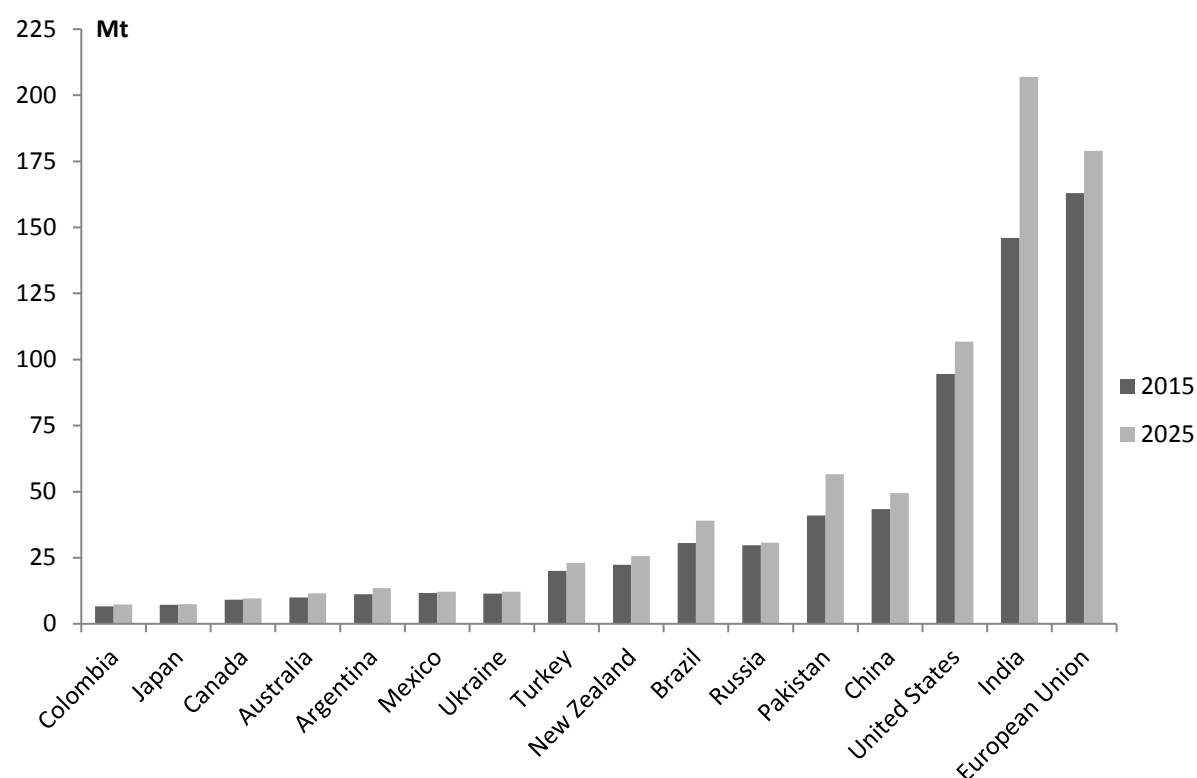
Figure 1.1 Milk production in 2015 and projections for 2025 (in million tonnes) for major producing countries and regions (Source: OECD, 2016).

Aforementioned benefits and environmental impacts associated with milk production vary largely across the world. One of the reasons for this variation is that dairy farming systems are heterogeneous. We find, for example, concentrate-based, high productive dairy systems in countries such as the Netherlands and the United States; grass-based dairy systems in countries such as Ireland and New Zealand; and smallholder dairy systems in countries such as Kenya and India. Even within one type of farming system, farms can vary largely and differ in biophysical factors, such as soil type and climatic conditions; in long term strategic decisions, such as the housing- or manure management system (Schröder et al., 2003; Nevens et al.,

2006); in short term tactical decisions, such as choice of feed ingredients or grazing regime; and in day-to-day management decisions, such as culling of cows or precision feeding (Nevens et al., 2006). All these differences can have an impact on the environmental performance of the dairy farming systems.

1.2 The conceptual framework

Multiple environmental impacts need to be addressed for benchmarking. Benchmarking dairy farming systems offers the opportunity to identify best farm practices and to provide guidance for improving farm performance (Stokes et al., 2007).

Benchmarking is defined by Camp et al. (1989) as "the search for those best practices that will lead to the superior performance" and, in this thesis, relates to the comparison of farms based on their environmental performance, in order to identify differences, and to deduce options to reduce environmental impacts. Figure 1.2 illustrates the conceptual framework for improving the environmental performance of dairy farming systems that was used in this thesis.

The framework for improving the environmental performance consists of two phases and 5 steps. It starts with the *benchmarking* phase, which includes three steps: (i) specify the goal and scope, (ii) select indicators, and (iii) assess and compare performances. The second phase is referred to as *evaluation and improvement*, and contains the remaining two steps: (iv) identify improvement options, and (v) implement and reflect on these options (adapted from van der Vorst et al. (2013) and de Olde et al. (2017)). The focus of this thesis is on the benchmarking phase, i.e. the first three steps in Figure 1.2. Below, I therefore describe each step in more detail, using this thesis's topic: benchmarking the environmental performance of dairy farming systems.
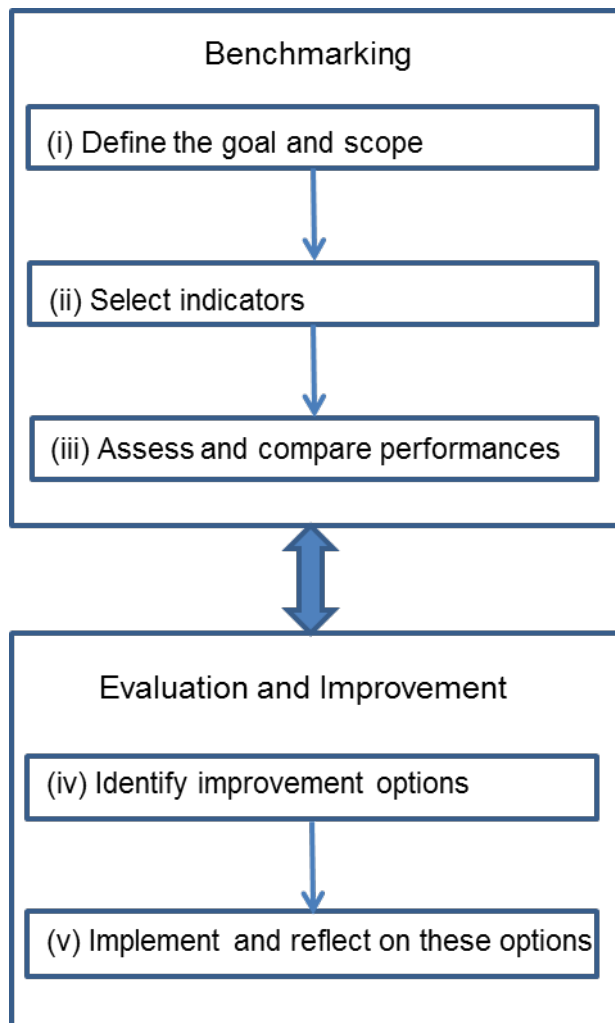
Figure 1.2 Conceptual framework for improving the environmental performance, adjusted from van der Vorst et al. (2013) and de Olde et al. (2017).

Step 1: In the first step, the goal and scope of the (benchmarking) assessment are defined, and the boundary of the system is determined. The goal is to reduce the environmental impact of dairy farming systems. The environmental impacts included (scope) are climate change, acidification, eutrophication, energy use and land use, being important impact categories in the dairy sector (Thomassen and De Boer, 2005; Yan et al., 2011). As the majority of these environmental impacts relate to the agricultural stage, including the production of purchased feed and processes on the dairy farm, the system boundary was defined from cradle-to-farm gate (see Figure 1.3). This system boundary includes processes involved up to the moment that milk leaves the farm gate, such as production of concentrates and roughage, fertilizers and keeping of the dairy herd (Schau and Fet, 2008).
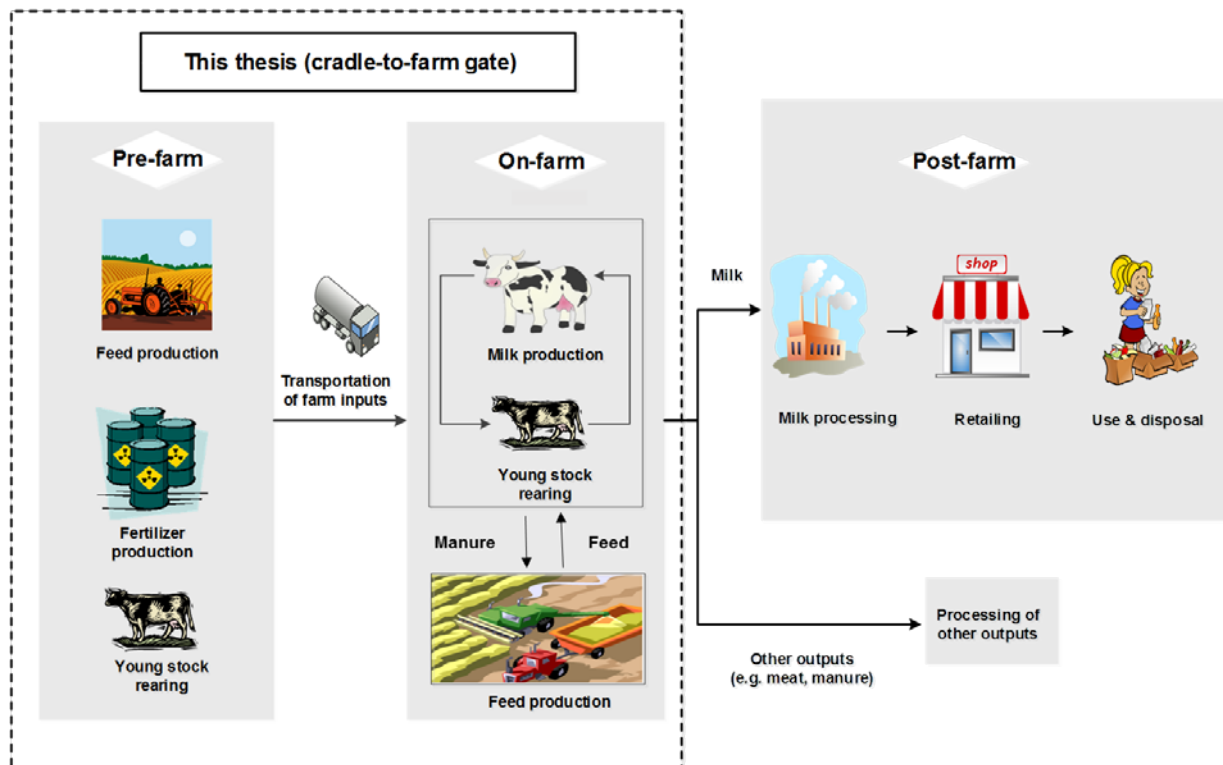
Figure 1.3 System boundary to assess environmental impacts of dairy farming systems in this thesis.

Step 2: The second step involves the selection of indicators to benchmark the farms. To assess the environmental performance of dairy farming systems, we need (a set of) environmental indicators. An indicator suitable for benchmarking should be relevant, measurable, valid, timely and understandable (Bell and Morse,1999; Lockie et al., 2002). To be more specific, *relevant* means that the indicator should be able to assess the major environmental performance of the farms; *measurable* means that the indicator can be quantified with the data that can be monitored and collected in a time and cost efficient way; *valid* means that the indicator should be measured both accurately and precisely. A measurement is accurate if the outcome is close to its true value (i.e. no bias or systemic error). A measurement is precise or has a high repeatability if you arrive at the same result in case you do the measurement several times; *timely* means that the indicator can provide early warning of potential environmental problems; *understandable* means that the results of the indicator can be interpreted easily by decision-makers, such as farmers and policy makers. In addition, to prevent biased conclusions and to account for potential trade-offs between environmental impacts, multiple environmental

6

indicators should be considered when benchmarking dairy farming systems. We therefore are in need of a set of indicators instead of a single indicator (Lyytimäki and Rosenström, 2008). The set of indicators should be able to reflect the current environmental state of the system comprehensively and reliably (Lebacq et al., 2013).

Potential environmental indicators to assess the environmental performance of dairy farming systems can be derived from several methods, including a nutrient balance approach, a material flow analysis or a life cycle assessment (Gerber et al., 2013). A description of the aforementioned methods are provided in the textbox below.

A nutrient balances (NB) approach computes the difference in nutrients entering and leaving a system, and allows computation of environmental indicators, such as nutrient use efficiency (NUE). NUE generally is defined as the amount of nutrients in valuable outputs of a system, such as milk, over the amount of nutrients in all inputs to that system, such as feed or fertilizer (Nevens et al., 2006). Additionally, an NB approach yields indicators such as the nutrient surplus of a system expressed per hectare of on-farm agricultural area or per kilogram valuable outputs. An NB approach generally is applied at farm level and 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 such as eutrophication (Oenema et al., 2003; Gourley et al., 2012).

A material flow analysis (MFA) uses the principle of mass balancing to quantify all in- and outflows of natural resources, energy and materials into, through and out of a given system. MFA can be further specified into bulk-material analysis which studies flows of bulk materials such as iron or plastic, and substance flow analysis which studies flows of group of elements such as nitrogen compounds or cadmium (Sonesson and Berlin, 2003). The output of MFA can help decision-makers identifying inefficient use of natural resources, energy and materials in the system. To gain more insight, it is often suggested to conduct MFA in combination with life cycle assessment to further evaluate the environmental impacts related to inefficiencies.

A 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 to the use of resources during the entire life of a product. 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). The mostly examined impact categories are: climate change, land use, energy use, acidification, eutrophication and the system boundary is often defined from cradle-to-farm gate (on-farm processes, e.g. manure management, milk and feed production, and pre-farm processes, e.g. fertilizer and feed production).

Step 3: The third step of the framework includes assessing and comparing the performance of systems, in this case, the environmental performance of dairy farming systems, using the indicators selected in step 2. To be able to compare systems, a reference value needs to be set. This reference value can be either a fixed value, e.g. an acceptable or desirable value, or a relative value, e.g. the average of the top 10% best performing systems (Diazabakana et al., 2014). Two cautions should be made here. First, the level of data uncertainty might affect the validity of benchmarking results. Required data contains uncertainty due to measurement or

observer errors (i.e. epistemic uncertainty), or the data can vary widely due to natural circumstances (e.g. impact of weather on grass yield). Second, systems should be compared against a desirable but achievable benchmark, e.g. best performance of a group of peers.

Step 4: The fourth step of the framework includes identifying improvement options. Identification of these improvement options can emerge from exploring the differences between, for example, least-scoring and best-scoring systems; or from alternative systems (i.e. different type of production systems) that have superior performances. An example of the latter could be the use of legumes instead of chemical fertilizers that is common in organic farms, but could improve the nutrient use efficiency of conventional farms as well. Improvement options can include incremental changes, but also adaptation of innovations. This thesis, however, does not focus on this fourth step.

Step 5: The fifth step of the framework includes implementing and reflecting on the identified improvement options. After a farmer has implemented the improvement option identified in step 4, for example, decision-makers could evaluate if the goal identified in step 1 has been reached. In addition, decision-makers can reflect upon the process and decisions that were made in all the steps to see what can be further improved in the future procedure for improving the environmental sustainability of the system. I do not focus on this step in this thesis.

1.3 Knowledge gaps

Most studies that address benchmarking of dairy farming systems focused on indicators of one environmental impact assessment method, either an NB approach (applied at the farm level) or LCA (applied at the (partial) chain level). Due to the simplicity of an NB approach and relatively low data requirement, indicators derived from this method have been used widely to assess the environmental performance of dairy farming systems (Oenema et al., 2001; Powell et al., 2010; Gourley, et al., 2012; Toma et al., 2013). In general, indicators from an NB approach are computed at farm level and do not include nutrient losses related to the production of farm inputs, such as purchased concentrates. Overlooking off-farm losses may

lead to biased benchmarking results. LCA can give a comprehensive evaluation of resource usage and emissions along the entire chain, but requires much more data. An NB approach at chain level that incorporates off-farm activities, therefore, might provide an opportunity to include off-farm processes while limiting data requirements. So far, however, no study has examined whether indicators derived from a chain-based NB approach can provide better insights than a farm-based NB when benchmarking the environmental performance of dairy farming systems.

Environmental indicators derived from an NB approach, moreover, generally do not specify the type of losses, nor the environmental impacts associated with those losses, such as the impact on acidification or climate change. Contrary to the indicators from an NB approach, indicators from an LCA do specify the types of losses, as well as the potential associated environmental impacts related to those losses. However, studies suggested that collection of data required to perform an LCA appears difficult and time consuming (Thomassen and De Boer, 2005). For benchmarking purposes, however, there is a need for a set of indicators, that are measurable, i.e. can be quantified regularly and, thus, are based on data that are technically and financially feasible to collect. Furthermore, these indicators should be relevant and timely, i.e. provide insight into the major environmental impacts of the dairy farming systems (Bélanger et al., 2012). Exploring correlations between various indicators can help to identify such a set of indicators (Lebacq et al., 2013). Previous studies have mainly focused on correlations between indicators within LCA and for non-dairy related products (e.g. meat, industrial materials). So far, no study examined correlations between environmental indicators from different environmental impact assessment methods within dairy production (Berger and Finkbeiner, 2011; Laurent et al., 2012; Röös et al., 2013). Therefore, studies on exploring correlations between NB and LCA indicators are needed, in order to identify an effective set of indicators that can be used as a proxy for the environmental performance of dairy farming systems.

Third, data collected for benchmarking dairy farming systems can by highly uncertain which can influence the validity of the indicators. This data uncertainty can be caused by the inherent

variations of the systems (i.e. natural variability), such as variation in grass productivity due to climate fluctuations; or by measurement errors and observational errors (i.e. epistemic uncertainty), such as measurement errors around the feed intake of the cows. Most studies that benchmark the environmental performance of dairy farming systems, however, do not consider the impact of data uncertainty (Schröder, 2003; Nevens, 2006; Powell, 2010). A few studies examined the impacts of epistemic uncertainty of nutrient flows on quantification of nutrient use efficiency (Mulier et al., 2003; Oenema et al., 2015) and the impacts of uncertainties of emission factors and natural variability of farm data on LCA results for assessing environmental performance of milk production (Basset-Mens et al., 2009; Chen et al., 2014). These studies, however, did not examine the impact of epistemic uncertainty on benchmarking results. Therefore, the effects of data uncertainty on benchmarking the performance of dairy farming systems should be explored.

Last but not least, to reduce environmental impacts while maintaining economic viability of dairy farming systems, we need to improve the eco-efficiency of the system. The concept of eco-efficiency was introduced in 1970s as the concept of "environmental efficiency" (Freeman et al., 1973). Later on, it was further developed to be the ratio between the economic value and environmental impact added by a system (Keffer and Shimp 1999; Seppälä et al., 2005). Although previous studies have already taken multiple environmental and economic indicators into account when benchmarking dairy farming systems (Fraser and Cordina, 1999; Stokes et al., 2007; Iribarren et al., 2011), they did not incorporate data uncertainty of the indicators into their analyses. As aforementioned, data uncertainty might cause biased benchmarking results. We, therefore, are in need of exploring a benchmarking technique that allows multiple indicators to be incorporated simultaneously, while considering the uncertainty associated with the indicators.

1.4 Objective and outline of the thesis

The objective of this thesis is to develop a sound method to benchmark the environmental performance of dairy farming systems. Acknowledging the variety of dairy farming systems in the world, this thesis focuses on specialized dairy farming systems in Europe. We defined specialized dairy farming systems as systems 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 first challenge in benchmarking the environmental performance of dairy farming systems is to select a set of indicators that can be used as a proxy for the environmental performance of those systems. Indicators from an NB approach and LCA are often applied in such a context and, therefore, will be evaluated to derive this set of indicators. Indicators from an NB approach, however, are often quantified at farm level, which may lead to biased benchmarking results because of neglecting losses related to off-farm processes, such as the production of purchased feed. In Chapter 2, indicators derived from a chain-based NB approach (i.e. including all on-farm processes as well as the production of purchased feed) are compared with indicators derived from a farm-based NB approach to explore the importance of benchmarking environmental performance of dairy farming systems at chain level.

Indicators derived from an NB approach are relatively easy to quantify, whereas the indicators from LCA provides more detailed insight into the type of losses and associated environmental impacts. In Chapter 3, we therefore, explore correlations between indicators derived from a NB approach and LCA, in order to identify an effective set of indicators that can be used as a proxy for the environmental performance of dairy farming systems.

The second challenge in benchmarking the environmental performance of dairy farming systems is to cope with data uncertainties. To meet this challenge, we start with exploring the impact of epistemic uncertainty of input parameters on benchmarking results for a relatively simple indicator: nutrient use efficiency, in Chapter 4. In addition, the epistemic uncertainty of input parameters that explain most of the output variance are identified and reduced, to illustrate how this improves benchmarking results.

To benchmark dairy farming systems comprehensively, multiple indicators need to be taken into account. In Chapter 5, the set of environmental indicators derived from Chapter 3 and an economic indicator are applied to assess the eco-efficiency of dairy farming systems, while accounting for data uncertainty during the benchmarking process.

In Chapter 6, the overall findings of this thesis are summarized and discussed. In addition, research implications are elaborated, and future research opportunities are pointed out.

1.5 References

Basset-Mens, C., Kelliher, F.M., Ledgard, S., Cox, N., 2009. Uncertainty of global warming potential for milk production on a New Zealand farm and implications for decision making. Int. J. Life. Cycle. Assess.14, 630-638.

Berger M., Finkbeiner, M., 2011. Correlation analysis of life cycle impact assessment, indicators measuring resource use. Int. J. Life Cycle Assess. 16, 74–81.

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.

Bell S, Morse S., 1999. Sustainability Indicators: Measuring the Immeasurable. Earthscan: London.

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, Goteborg, Sweden.

Camp, R.C., 1989. Benchmarking: The Search for Industry Best Practices that Lead to Superior Performance, ASQC Quality Press, Milwaukee, WI.

Chen, X., Corson, M.S., 2014. Influence of emission-factor uncertainty and farm-characteristic variability in LCA estimates of environmental impacts of French dairy farms. J. Clean. Prod. 81, 150-157.

Diazabakana, A., Latruffe, L., Bockstaller, C., Desjeux, J, Finn, J., Kelly, E, Ryan, M. and Uthes, S., 2014. A Review of Farm Level Indicators of Sustainability with a Focus on CAP and FADN. FLINT. <http://www3.lei.wur.nl/flint/downloads/reports/FLINT%20WP120.D1>.

De Olde, E.M., Bokkers, E.A.M., De Boer, I.J.M., 2017. The Choice of the Sustainability Assessment Tool Matters: Differences in Thematic Scope and Assessment Results. Ecol. Econ. 136, 77-85.

De Vries, M., De Boer, I.J., 2010. Comparing environmental impacts for livestock products: A review of life cycle assessments. Livest. Sci. 128, 1-11.

Dolman, M.A., Sonneveld, M.P.W., Mollenhorst, H., De Boer, I.J.M., 2014. Benchmarking the economic, environmental and societal performance of Dutch dairy farms aiming at internal recycling of nutrients. J. Clean. Prod. 73, 245–252.

FAO, 2013. Muehlhoff, E., Bennett, A., McMahon, D. Milk and Dairy Products in Human Nutrition (2013), by, Food and Agriculture Organisation of the United Nations (FAO), Rome. E-ISBN: 978-92-5-107864-8 (PDF).

Freeman, M.A., Haveman, K., Kneese, A.V., 1973. The Economics of Environmental Policy. John Wiley and Sons, New York.

Fraser, I., Cordina, D., 1999. An application of data envelopment analysis to irrigated dairy farms in Northern Victoria, Australia. Agric. Syst. 59, 267-282.

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.I., 2012. Farm-scale nitrogen, phosphorus, potassium

and sulphur balances and use efficiencies on Australian dairy farms. Anim. Prod. Sci. 52, 929–944.

Gerber, P.J., Steinfeld, H., Henderson, B., Mottet, A., Opio, C., Dijkman, J., Falcucci, 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.

ISO, I., 1997. 14040: Environmental Management: Life Cycle Assessment: Principles and Framework. European Committee for Standardization (CEN), Brussels, Belgium.

ISO, I., 1998. 14041: Environmental management-life cycle assessment-goal and scope definition and inventory analysis. European Committee for Standardization (CEN), Brussels, Belgium.

ISO, I., 2000. 14042: environmental management—life cycle assessment—life cycle impact assessment. Geneva, Switzerland. European Committee for Standardization (CEN), Brussels, Belgium.

ISO, I., 2000. 14043: Environmental Management–Life Cycle Assessment–Life Cycle Interpretation. European Committee for Standardization (CEN), Brussels, Belgium.

Iribarren, D., Hospido, A., Moreira, M.T. Feijoo, G., 2011. Benchmarking environmental and operational parameters through eco-efficiency criteria for dairy farms. Sci. Total Environ., 409 ,1786–1798.

Keffer, C., Shimp, D., 1999. Eco-efficiency Indicators and Reporting. World Business Council for Sustainable Development (WBCSD), London.

Leip, A., Billen, G., Garnier, J., Grizzetti, B., Lassaletta, L., Reis, S., Simpson, D., Sutton, M.A., De Vries, W., Weiss, F., 2015. Impacts of European livestock production: nitrogen, sulphur, phosphorus and greenhouse gas emissions, land-use, water eutrophication and biodiversity. Environ. Res. Lett. 10, 115004.

Laurent, A., Olsen, S.I., Hauschild, M.Z., 2012. Limitations of carbon footprint as indicator of environmental sustainability. Environ. Sci. Technol. 46,4100–4108.

Lyytimäki, J., Rosenström, U., 2008. Skeletons out of the closet: effectiveness of conceptual frameworks for communicating sustainable development indicators. Sustain. Dev.16, 301-313.

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.

Lebacq, T., Baret, P.V., Stilmant, D., 2013. Sustainability indicators for livestock farming. A review. Agron. Sustain. Dev., 33, 311-327.

Mugera, A.W., 2013. Measuring technical efficiency of dairy farms with imprecise data: a fuzzy data envelopment analysis approach. Aust. J. Agr. Resour. Econ. 57, 501-520.

Mulier, A. , Hofman, G. , Baecke, E. , Carlier, L. , De Brabander, D. , De Groote, G. , 2003. A methodology for the calculation of farm level nitrogen and phosphorus balances in Flemish agriculture. Eur. J. Agron, 20, 45–51.

Nevens, F., Verbruggen, I. , Reheul, D., Hofman, G., 2006. Farm gate nitrogen surpluses and nitrogen use efficiency of specialized dairy farms in Flanders: Evolution and future goals. Agric. Syst. 88, 142–155.

OECD,2016. OECD-FAO Agricultural Outlook 2016-2025, OECD Publishing, Paris. DOI: http://dx.doi.org/10.1787/agr_outlook-2016-en

Oenema, J., Koskamp, G.J., Galama, P.J., 2001. Guiding commercial pilot farms to bridge the gap between experimental and commercial dairy farms: the project 'Cows & Opportunities'. Neth. J. Agr. Sci. 49, 277-296.

Oenema, J., Burgers, S., van Keulen, H., van Ittersum, M., 2015. Stochastic uncertainty and sensitivities of nitrogen flows on dairy farms in The Netherlands. Agric. Syst. 137, 126-138.

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.

Powell, J.M., Gourley, C.J.P., Rotz, C.A., Weaver, D.M., 2010. Nitrogen use efficiency: a potential performance indicator for dairy farms. Environ. Sci. Policy. 13, 217–228.

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.

Schau, E.M., Fet, A.M., 2008. LCA studies of food products as background for environmental product declarations. Int. J. LCA. 13, 255-264.

Steinfeld, H., Gerber, P., Wassenaar, T., Castel, V., Rosales, M., de Haan, C., 2006. Livestock's long shadow. Food And Agriculture Organization of the United Nations.

Schröder, J.J., Aarts, H.F.M., Ten Berge, H.F.M., Van Keulen, H., Neeteson, J.J., 2003. An evaluation of whole-farm nitrogen balances and related indices for efficient nitrogen use. Eur. J. Agron. 20,33–44.

Seppälä, J., Melanen, M., Mäenpää, I., Koskela, S., Tenhunen, J., Hiltunen, M.-R., 2005. How can the eco-efficiency of a region be measured and monitored? J. Ind. Ecol. 9,117–130.
Stokes, J., Tozer, P., Hyde, J., 2007. Identifying efficient dairy producers using data envelopment analysis. J. Dairy Sci. 90, 2555-2562.

Sefeedpari, P., Rafiee, S., Akram, A., 2012. Selecting energy efficient poultry egg producers: a fuzzy data envelopment analysis approach. Int. J. Appl. Oper. Res. 2, 77-88.

Sonesson, U., Berlin, J., 2003. Environmental impact of future milk supply chains in Sweden: a scenario study. J. Clean. Prod. 11, 253-266.

Toma, L., March, M., Stott, A.W., Roberts, D.J., 2013. Environmental efficiency of alternative dairy systems: A productive efficiency approach. J. Dairy. Sci. 96, 7014-4031.

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.

Van der Vorst, J.G., Peeters, L., Bloemhof, J.M., 2013. Sustainability Assessment Framework for Food Supply Chain Logistics: Empirical Findings from Dutch Food Industry. Int. J. Food. Sys. Dynam., 130-139.

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.

Chapter 2

Nutrient balance at chain level: a valuable approach to benchmark nutrient losses of milk production systems?

W. Mu[1,2], C.E. van Middelaar[1], J.M. Bloemhof[2], J. Oenema[3], I.J.M. de Boer[1]

[1] Animal Production Systems group, Wageningen University, P.O. Box 338, 6700 AH Wageningen, the Netherlands

[2] Operations Research and Logistics group, Wageningen University, P.O. Box 8130, 6700 EW Wageningen, the Netherlands

[3] Plant Research International, Wageningen University, P.O. Box 616, 6700 AP Wageningen, the Netherlands

Abstract

A nutrient balance approach is often used to quantify losses of nutrients, such as nitrogen and phosphorus, that contribute to environmental problems such as eutrophication. A nutrient balance generally is computed at farm level, implying that nutrient losses related to pre-farm processes, such as production of purchased feed, are neglected. Using a nutrient balance at farm level to benchmark livestock systems or individual farms that differ in, for example, amount of purchased concentrates, however, may lead to biased conclusions. To determine whether a nutrient balance that accounts for losses during production of purchased feed (i.e. a chain balance from cradle-to-farm-gate) are more suited to benchmark nutrient losses of milk production systems or individual farms than a nutrient balance at farm level, we analysed 19 Irish grass-based dairy farms and 13 Dutch concentrate-based dairy farms. For each farm, we computed nitrogen and phosphorus losses at farm and chain level (i.e. from cradle-to-farm-gate, including losses from purchased feed production), and expressed these losses per ton fat-and-protein-corrected milk. An independent T-test and Wilcoxon-Mann-Whitney test were used to examine the difference in mean losses at farm and chain level between Irish and Dutch systems. Regression analysis was used to determine if the ranking of the farms changes using a nutrient balance at a farm or chain level. Results show that on average, Irish farms had higher nitrogen losses per ton milk than Dutch farms, both at farm (Irish=20; Dutch=8 in kg N/ton milk) and chain level (Irish=22; Dutch=11 in kg N/ton milk). Phosphorus losses per ton milk, on the other hand, did not differ between Irish and Dutch farms at farm (Irish=0.3; Dutch=0.1 in kg P/ton milk) or chain level (Irish=0.8; Dutch=1.0 in kg P/ton milk). Regression analysis revealed that the nutrient balance at chain level could be accurately predicted from the nutrient balance at farm level (R^2=0.992 for N; R^2=0.910 for P); whereas in case of phosphorus, the slope tended to differ between Irish and Dutch farms ($p<0.10$). Ranking 32 farms based on the nitrogen balance at farm or chain level, therefore, showed a similar pattern, whereas the ranking pattern based on the phosphorus farm balance differed from the pattern based on the chain balance. We concluded, therefore, that to benchmark nutrient losses of dairy systems, a nutrient balance at farm level can be used if differences in on-farm losses

between systems are large, and pre-farm losses related to, e.g. production of purchased concentrates, are relatively unimportant. To benchmark individual farms, a nutrient balance at farm level can be used only if changes in pre-farm losses per unit change in on-farm losses are similar across farms. A chain level balance of a sample set, however, is required to verify these conditions.

2.1 Introduction

The world population is projected to increase from about 7 billion in 2013 to about 10 billion in 2050 (UN report, 2013). Population growth in combination with an improvement of living standards is expected to increase the demand for highly nutritious products, especially animal-source food, such as milk (FAO, 2006). The Food and Agriculture Organization of the United Nations (FAO) predicts global consumption of meat and dairy products to increase by 82% between 2000 and 2050, which implies an additional production of 466 million tonnes of milk (Boland et al., 2013).

Sustainable intensification of milk production may provide a possibility to meet this growing demand for animal-source food. Sustainable intensification implies increasing production levels from existing land while reducing the pressure on the environment (Garnett et al., 2014; Mont et al., 2014). Inefficient use of the nutrients nitrogen (N) and phosphorus (P) in agriculture is a main cause of environmental pressure, and might cause problems such as eutrophication, acidification and global warming (Volk et al., 2009; Djekic et al., 2014), but also soil degradation (Sutton et al., 2013).

One common method to quantify the environmental impact of nutrient use is the nutrient balance (NB) approach (Oenema et al., 2003). An NB computes the difference in nutrients entering and leaving a system, and allows computation of environmental indicators, such as nutrient use efficiency (NUE). NUE generally is defined as the amount of nutrients in valuable outputs of a system over the amount of nutrients in all inputs of that system (Nevens et al., 2006). Additionally, an NB yields indicators such as the nutrient surplus of a system expressed per hectare of on-farm agricultural area or per kilogram valuable outputs. An NB of a dairy farming system generally is computed at farm level. The balance of N and P are mostly studied because N and P are two major nutrients that can limit crop growth and losses of N and P can cause environmental problems (Thomassen and De Boer, 2005; Huhtanen et al., 2011; Toma et al., 2013; Dolman et al., 2014). An NB approach at farm level does not incorporate detailed

information on material and nutrient flows within a farm, but only quantifies the inputs and outputs of the farm; therefore, it requires relative little data. An NB approach has been used as a valuable approach to provide farmers with farm-specific advice (i.e. identify best practices) (Oenema et al., 2001) and to benchmark environmental impacts of livestock systems (Halberg, 1999; Bengtsson et al., 2003; Powell et al., 2010; Gourley, et al., 2012; Toma et al., 2013).

Indicators derived from an NB approach at farm level, however, do not include nutrient losses related to the production of farm inputs, such as purchased concentrates. A farm-specific advice directed at reducing nutrient losses at the farm might affect the type and amount of purchased inputs, and, therefore, nutrient losses related to the production of these inputs. Similarly, comparing the environmental performance of contrasting farming systems based on a farm-based NB may lead to biased results because of differences in farm inputs. Concentrates-based dairy systems, for example, generally purchase large amounts of feed, whereas grass-based dairy systems use mainly on-farm produced feed. Comparing both systems using indicators derived from a farm-based NB may lead to biased conclusions in favour of concentrates-based dairy systems.

An NB that incorporates off-farm activities, such as production of purchased feed, therefore, might be needed to make a fair comparison (Gerber et al., 2014). Although life cycle assessment (LCA) can give a comprehensive evaluation of resource usage and emissions along the entire chain, it appears difficult and time consuming to collect all data required for LCA indicators (Thomassen and De Boer, 2005). The objective of this paper is to examine whether indicators derived from a chain-based (i.e. from cradle-to-farm-gate) NB can provide better insights than a farm-based NB when benchmarking different dairy systems and individual farms. We, therefore, compared indicators from an NB at farm and chain level for 19 grass-based dairy farms (Ireland) and 13 concentrates-based dairy farms (The Netherlands).

2.2 Material and methods

2.2.1 Data

To assess NB indicators of contrasting milk production systems, we used data of farms from Dairyman. Dairyman was a project in the INTERREGIVB program co-funded by the European Regional Development Fund, which aimed to improve regional prosperity through better resource utilization on 113 dairy farms in different European countries and stakeholder cooperation (Dairyman, 2010). The database contains detailed information on farm characteristics, such as production parameters (e.g. milk yield per cow, replacement rate) and type and amount of purchased production inputs, including the nutritional value of feed inputs.

We identified 32 specialised dairy farms from Dairyman and determined indicators for 2010, i.e. 19 farms are from Ireland (grass-based) and 13 farms are from the Netherlands (concentrates-based; see Table 2.1). We defined specialized 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. As indicated in Table 2.1, Irish farms use more than 80% of their farm area and Dutch farms use more than 70% of their farm area to grow grass. This is in line with the former derogation regulation of the European nitrate directive in 2010 that prescribed that farms with at least 70% (NL) or 80% (IR) grassland can apply maximally 250 kg manure-N per ha per year instead of 170 kg manure-N per ha per year (EU, 2010; European Communities, 2010). In general, compared to the Irish situation, Dutch farms have a higher stocking rate, feed more concentrates per cow, purchase more roughages per cow, and have a higher milk production per cow and per ha of land.

Table 2.1 Farm characteristics of specialized dairy farms in Ireland (IR) and the Netherlands (NL) used in this study

Characteristics*	IR (n**=19) Mean(SD)	NL (n=13) Mean(SD)
Agricultural area (ha)	72(25.1)	53(18.1)
-- Grassland area	61 (22.8)	42(12.7)
-- Maize land area	1(2.0)	11(7.6)
Number of dairy cows***	109(29.7)	122(32.8)
Milk production per cow (kg cow^{-1} year^{-1})	5304(646.7)	8732(776.3)
Milk production per ha (kg ha^{-1} year^{-1})	8341(1855.6)	20875(5675.0)
Stocking rate (dairy cows ha^{-1}year^{-1})	1(0.5)	3(0.7)
Use purchased concentrates per cow (kg cow^{-1} year^{-1})	814(180.6)	2053(249.4)
Purchased concentrate (ton year^{-1})	89(32.6)	253(68.1)
Purchased roughage (ton DM year^{-1})	62(68.4)	194(108.9)

*According to the Wilcoxon-Mann-Whitney test, all characteristics are significantly different (P<0.05) between NL and IR,
except for the number of dairy cows and live weight of the cows
** n represent the number of farms included in this study
*** Dutch farms only have Holstein Friesians, whereas Irish farms use different types of breed, including Holstein Friesian, Swedish Red, Norwegian Red, Jersey, and others.

2.2.2 System boundaries

We compared the impact of calculating an NB from cradle-to-farm-gate (chain) or at the farm level, as illustrated in Figure 2.1. Inputs included in the NB at farm level were concentrates, roughage, organic fertilizer, mineral fertilizers and atmospheric deposition for N. Outputs included were milk, animals, and crops. The farm itself was considered as a black box. We calculated net inputs or net outputs of products that were both purchased and sold, such as animals. In case the animals output was larger than the animal input, the difference between input and output was considered a net output, whereas in case the animal output was smaller than the animal input, the difference between in- and output was considered as a net input (Godinot et al., 2014). For the case of manure, we subtract manure output from the organic fertilizer input, so when manure output exceeds organic fertilizer input, it was shown as a negative net input instead of a positive net output. Calculating a net in- or output for animals, crops and manure does not affect the NB at farm level. At chain level, however, estimating nutrient losses related to production of animals and manure entering the farm is avoided. Stock

changes (defined as final stock of the year 2010 – initial stock of the year 2010) of the concentrates, roughages, animals and fertilizers were included in the computation of NB indicators.
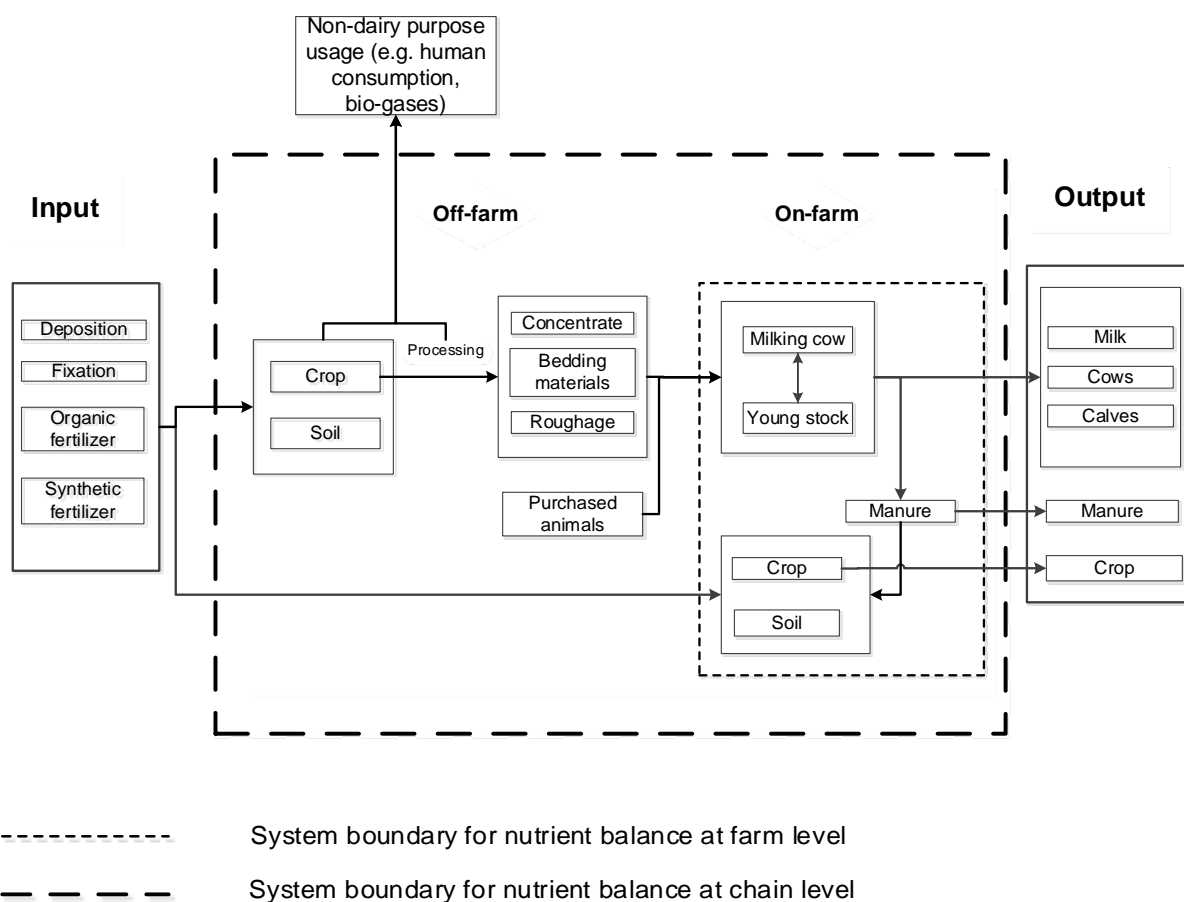
Figure 2.1 System boundaries for nutrient balance at farm and chain level

At chain level, the system boundary included all relevant processes from cradle-to-farm-gate, i.e. all on-farm processes as well as the production of purchased feed. The chain level did not include processes after the farm gate (i.e. milk processing, retailing, consumer consumption and disposal stages), because these processes were assumed to be similar for Ireland and the Netherlands, and, furthermore, cannot be influenced by farmer's decisions. Nutrient losses from on-farm processes were adopted from the NB at farm level. For production of purchased feed, we considered nutrient losses during cultivation (including 5% storage losses). Nutrient losses during processing of feed ingredients were assumed to be negligible (Sterk, 2014). Nutrient losses during production of mineral fertilizers were not included in the analysis

because the contribution of these losses to the total surplus was found to be minor (Godinot et al., 2014).

2.2.3 Calculation of nutrient losses

First, we computed the nutrient losses of N and P for each of the 32 dairy farms at farm level using the following formulas:

$$B^{\text{farm}}_{\text{nutrient}} = (I1 - S1 + I3 - S3 + I4 - S4 + I5 - O3 - S5 + I6) - (O1 + O2 - I2 + S2 + O4)$$

$$L^{\text{farm}}_{\text{nutrient}} = \text{Max}(B^{\text{farm}}_{\text{nutrient}}, 0)$$

Where, $B^{farm}_{nutrient}$ is the NB at farm level; I refers to the inputs of N or P via concentrates (*I1*), animals (*I2*), roughage (*I3*), mineral fertilizer (*I4*), organic fertilizer (*I5*) and atmospheric deposition (*I6*); O refers to the outputs of N and P via milk (*O1*), animals (*O2*), organic fertilizer (*O3*) and plant products (*O4*); and S refers to stock changes expressed in N or P of concentrates (S1), animals (*S2*), roughages (*S3*), mineral fertilizer (*S4*) and organic fertilizer (*S5*). Table 2.2 shows relevant inputs, outputs, calculation processes and references used. We do not consider N fixed by clover as an input because of the low prevalence of clover on the farms and, subsequently, a small contribution to the total N input (Gourley et al., 2007). Phosphorus deposition was not included either, because it was regarded as insignificant (Steinshamn et al., 2004).

Table 2.2 Nutrient inputs and outputs included in the nutrient balance at farm and chain level

Elements	Farm	Off-farm	Chain	Calculation[3] and References
Input:				
Purchased concentrates (I1)	√		√	Q x nutrient content of concentrates (Dairyman[4], 2010)
Purchased animals (I2)	√		√	Q x nutrient content of animals (Raison et al., 2006)
Purchased roughages[1] (I3)	√		√	Q x nutrient content of roughage (Dairyman, 2010)
Purchased mineral fertilizer (I4)	√		√	Q x nutrient content of mineral fertilizer (Dairyman, 2010)
Purchased organic fertilizer (I5)	√		√	Q x nutrient content of organic fertilizer (Dairyman, 2010)
Atmospheric deposition[2] (I6)	√		√	Average value of relevant region (EMEP, 2007)
N&P input for producing concentrates[3] (I7)		√	√	Q x nutrient input for producing concentrates (Vellinga et al., 2013)
N&P input for producing roughages[3] (I8)		√	√	Q x nutrient input for producing roughages (Vellinga et al., 2013)
N&P input for producing by-products[3]		√	√	Assumed to be 0, because the economic allocation of by-products is 0
Stock change[5]:				
Stock change of concentrates (S1)	√		√	Q x nutrient content of concentrates (Dairyman, 2010)
Stock change of animals (S2)	√		√	Q x nutrient content of animals (Dairyman, 2010)
Stock change of roughages (S3)	√		√	Q x nutrient content of roughages (Dairyman, 2010)
Stock change of mineral fertilizer (S4)	√		√	Q x nutrient content of mineral fertilizer (Dairyman, 2010)
Stock change of organic fertilizer (S5)	√		√	Q x nutrient content of organic fertilizer (Dairyman, 2010)
Output:			√	
Milk output (O1)	√		√	Q x nutrient content of milk (Dairyman, 2010)
Animal output (O2)	√		√	Q x nutrient content of animal (Raison et al., 2006)
Organic fertilizer output(O3)	√		√	Q x nutrient content of organic fertilizer (Dairyman, 2010)
Plant output (O4)	√		√	Q x nutrient content of plant (Raison et al., 2006)
Off-farm plant outputs for producing concentrates		√	√	Q* x nutrient content of plant (CVB,2008)
(O5)		√	√	Q* x nutrient content of plant (CVB,2008)
Off-farm plant outputs for producing roughages (O6)				

[1] We selected five most commonly used roughages in the dataset of Dairyman, i.e. grass silage, maize silage, hay, wheat straw and alfalfa.
[2] Atmosphere deposition is only applicable for nitrogen; for phosphorus the amount of deposition is small and, therefore, neglected.
[3] Q = actual quantity of product purchased, sold or stock change, obtained from Dairyman for year 2010 (in kg or numbers)
  Q*=% of certain concentrate ingredients e.g. soybean, rapeseed (based on the estimated feed composition from Dairyman) x consumed quantity of concentrate or roughage of certain farm in 2010
[4] Dairyman project collected data from farms based on farm records and by interviewing farmers.
[5] Stock change was calculated as final stock of the year 2010 - initial stock of the year 2010. Therefore it can be an inflow (if S is positive) or an outflow (if S is negative).

To compute nutrient losses at farm level ($L_{nutrient}^{farm}$), we took the maximum value of the NB ($B_{nutrient}^{farm}$) and 0. A positive P balance implies that P is lost to the environment through leaching and runoff, contributing to increased eutrophication of ground and surface water, and will be referred to as P loss. A positive N balance implies that N is lost to the environment through, e.g., volatilisation of ammonia or nitrous oxide, or through runoff and leaching of dissolved nitrate. Different losses of N are thus summed into one indicator, and will be referred to as N loss. For many years, there has been a supply of excessive P to the field, resulting in increased levels of soil P. As a result, farmers decreased or even stopped the application of mineral fertilizer P. When we consider the NB on a yearly basis, in some cases this decline in P application led to a negative P balance. Because this negative P balance does not imply a loss, nutrient losses were set to zero.

Farm nutrient losses ($L_{nutrient}^{farm}$) were expressed per ha of farm land or per ton of fat-and-protein-corrected milk (FPCM), i.e. milk corrected to a fat percentage of 4.0% and a protein content of 3.3% using the following formula (Product Board Animal Feed 2008):

FPCM (ton) = Milk (ton) x [0.337 + 0.116 x Fat (%) + 0.06 x Protein (%)]

Nutrient losses per ha at farm level can reflect the local or regional impacts of dairy farming, whereas nutrient losses per unit of product provide insights into the efficiency of a production system (Halberg et al., 2005).To express nutrient losses per ton FPCM, $L_{nutrient}^{farm}$ were allocated to the various outputs based on their relative economic value (i.e. economic allocation). Economic allocation is a commonly used allocation method in LCA studies on agricultural products (de Vries and de Boer, 2010).

Second, in order to calculate chain losses ($L_{nutrient}^{chain}$), we extended the farm NB by including off-farm NB ($B_{nutrient}^{off-farm}$) related to cultivation of purchased feed using the following formulas:

$$B_{nutrient}^{off-farm} = (I7 + I8) - (O5 + O6)$$

$$L_{nutrient}^{off-farm} = Max(B_{nutrient}^{off-farm}, 0)$$

$$L_{nutrient}^{chain} = L_{nutrient}^{farm} + L_{nutrient}^{off-farm}$$

Where, I refers to the input of N or P required for production of consumed concentrates (*I7*) and roughage (*I8*) purchased by the farm and O refers to the output of N or P via harvested crops which are used for producing consumed concentrates (*O5*) and roughage (*O6*) that are purchased by the farm. $L_{nutrient}^{off-farm}$ represents losses during the cultivation of purchased feed products. To compute $L_{nutrient}^{off-farm}$, we took the maximum value of NB ($B_{nutrient}^{off-farm}$) and 0. In this way, a negative NB is not treated as a loss. Soybean production in Argentina, for example, was found to have a negative N balance. Nitrogen losses, in this case, were assumed to be zero. Nutrient losses at the chain level ($L_{nutrient}^{chain}$) consist of nutrient losses on-farm ($L_{nutrient}^{farm}$) and nutrient losses off-farm ($L_{nutrient}^{off-farm}$). Similar to calculations at farm level, nutrient losses of a multifunctional process were allocated to the various outputs based on economic allocation. The economic value of by-products (e.g. fresh beet pulp) was assumed to be 0 (Vellinga et al., 2013).

$L_{nutrient}^{chain}$ was only expressed per ton FPCM because averaging the losses over the land area across the chain without detailed information on the local environmental pressure does not provide insights into the local or regional impact of a production system. This aspect will be further elaborated in the discussion section.

2.2.4 Statistical analysis

In order to explore whether nutrient losses at chain level provide different insights into the comparison between grass based (IR) and concentrate based (NL) systems than nutrient losses at farm level, we compared the mean difference in N and P losses between Irish and Dutch farms, at farm and chain level. Normality of data was checked with Shapiro-Wilk test (P >0.05). For normally distributed data, mean N and P losses were compared using the

independent T-test (P<0.05), whereas for non-normally distributed data, the Wilcoxon-Mann-Whitney test (P<0.05) was used.

To determine if the ranking of the individual farms across systems changes based on an NB at farm or chain level, we used the following linear regression model:

$$Y_i = \beta_0 + \beta_1 X_1 + \beta_2 X_{2i} + \beta_3 X_1 * X_{2i} + \varepsilon_i$$

where, Y_i is chain level nutrient losses per ton FPCM for farm i, X_1 is a binary variable (0 for Irish farms, 1 for Dutch farms), X_{2i} is farm level nutrient losses per ton FPCM for farm i, β_0 is the intercept for Irish farms, $\beta_0 + \beta_1$ is the intercept for Dutch farms, β_1 shows the difference in intercept between IR and NL (i.e. difference in the absolute level of off-farm losses); β_2 is the slope for Irish farms, $\beta_2 + \beta_3$ is the slope for Dutch farms, β_3 shows the difference in slope between IR and NL (i.e. difference in the change in off-farm losses per unit change in on-farm losses). In case β_3 is 0, ranking of farms based on the farm balance tend to be in line with the ranking based on chain balance.

All statistical analyses in this study were performed using the software SPSS (SPSS, 2011).

2.3 Results

Nitrogen losses at farm and chain level

Table 2.3 shows the mean and standard deviation of N inputs, outputs and losses of Irish and Dutch dairy farms. Results are expressed in kg N per ha and per ton FPCM. Major N inputs on Irish farms are mineral fertilizers (199.1 kg N/ha; 21.7 kg N/ton FPCM), whereas for Dutch farms major N inputs are concentrates (174.6 kg N/ha; 7.7 kg N/ton FPCM) and mineral fertilizers (117.8 kg N/ha; 5.3kg N/ton FPCM). Irish farms do not import or export any organic fertilizer, therefore the N input via organic fertilizer is 0. On average, the N input via organic fertilizer on Dutch farms is negative (-78.2 kg N/ha; -3.1 kg N/ton FPCM), which means that the N output of organic fertilizers (i.e. manure) exceeds the input.

Table 2.3 Mean and standard deviation (SD) of nitrogen (N) inputs, outputs and losses at farm level, of Irish and Dutch dairy farms, expressed in kg N per ha and per ton FPCM

Indicators	Per ha		Per ton FPCM	
	IR(n=19) Mean(SD)	NL (n=13) Mean(SD)	IR (n=19) Mean(SD)	NL (n=13) Mean(SD)
N inputs[*]				
Mineral fertilizer	199.1(55.9)	117.8(53.7)	21.7(5.2)	5.3(2.1)
Organic fertilizer	0.0(0.0)	-78.2(69.6)	0.0(0.0)	-3.1(2.1)
Roughage	12.4(17.1)	82.8(64.8)	1.2(1.3)	3.7(2.4)
Concentrates	40.1(18.4)	174.6(69.2)	4.4 (2.0)	7.7(1.4)
Atmospheric deposition	8.1(0.8)	36.0(0.2)	0.9(0.2)	1.7(0.5)
Total inputs	259.7(64.4)	333.0(80.0)	28.1(4.8)	15.2(2.3)
N outputs[*]				
Milk	46.7(10.5)	120.3(32.6)	5.0(0.1)	5.4(0.1)
Animal	11.4(4.5)	13.5(5.7)	1.3(0.6)	0.6(0.2)
Plants	0.0(0.0)	2.1(6.2)	0.0(0.0)	0.2(0.4)
Total outputs	58.1(11.9)	135.9(34.5)	6.3(0.6)	6.2(0.4)
N losses	217(61.3)	196(56.5)	20(4.3)	8(1.6)

* All the N inputs and outputs are net inputs and outputs, which takes consideration of stock change. Organic fertilizer inputs = purchased organic fertilizer - stock change - organic fertilizer outputs. Animal outputs = animal sold-animal input + stock change

When expressed per ha, total N inputs of the Irish farms (259.7 kg N/ha) are more than 20% lower than the total inputs of the Dutch farms (333 kg N/ha). Milk is the main N output for both Irish and Dutch farms. Total N outputs of the Irish farms (58.1 kg N/ha) are less than half the total outputs of the Dutch farms (135.9 kg N/ha). Farm level N losses per ha on Irish farms (217 kg N/ha on average) is not statistically different (p=0.34) from the N losses per ha on Dutch farms (196 kg N/ha). When expressed per ton FPCM, total N inputs of the Irish farms (28.1 kg N/ton FPCM) are almost twice as high as of the Dutch farms (15.2 kg N/ton FPCM). Total N outputs, however, are about the same for both farms (6.3 kg N/ton FPCM for the Irish and 6.2 kg N/ton FPCM for the Dutch farms). When expressed per ton FPCM, N losses are significantly higher on Irish farms than on Dutch farms at both farm and chain level (Table 2.4).

Table 2.4 Mean, standard deviation (SD) and mean difference significance level (p) of nitrogen (N) losses at farm and chain level of the Irish and Dutch dairy farms, expressed per ton FPCM.

Indicators[*]	Per ton FPCM		
	IR Mean(SD)	NL Mean(SD)	Mean difference (p)
N losses (kg N)			
On-farm	20(4.3)	8(1.6)	12(0.00)
Off-farm	2(0.9)	3(0.5)	1(0.00)
Chain	22(4.5)	11(1.4)	11(0.00)

[*]According to the results of Shapiro-Wilk test for normality, all indicators related to nitrogen for both Irish farms and Dutch farms are normally distributed (P>0.05).

Production of concentrate ingredients contributed most to the N losses from off-farm activities in case of both Irish (76%) and Dutch farms (79%). For Irish farms, cultivation of wheat straw and grass silage contributed most to N losses related to production of roughage, while cultivation of cereals and other concentrates (i.e. the average of soybean meal, rapeseed meal and cereal) contributed most to N losses related to production of concentrates. For Dutch farms, cultivation of maize silage and grass silage were the main contributors to the losses related to roughage production, and cultivation of rapeseed meal and cereal to the losses of concentrate production.

Irish farms (2 kg N/ton FPCM), therefore, have a lower off-farm N loss than Dutch farms (3 kg N/ton FPCM), because Dutch farms rely more on purchased feed products to produce milk than Irish farms. Table 2.5 shows that off-farm N losses per ton FPCM are lower for Irish farms than for Dutch farms ($\beta_1 = 2.83$, p <0.001), but the change in off-farm losses per unit change in on-farm losses is similar on both type of farms ($\beta_3 = -0.17$, p =0.15). As a result, adding the off-farm losses to the on-farm losses does not change the order of the farms when ranking them based on N losses. Figure 2.2 shows the ranking of the farm based on N losses at either farm and chain level. Farms are ranked based on N losses at farm level. Results show a clear cluster of Irish farms and Dutch farms, i.e. all the Dutch farms are ranking ahead of Irish farms. Furthermore, the ranking of the 32 farms based on the N losses at either farm or chain level showed a similar pattern.

Table 2.5 Results for regression of nitrogen losses at farm level on nitrogen losses at chain level.

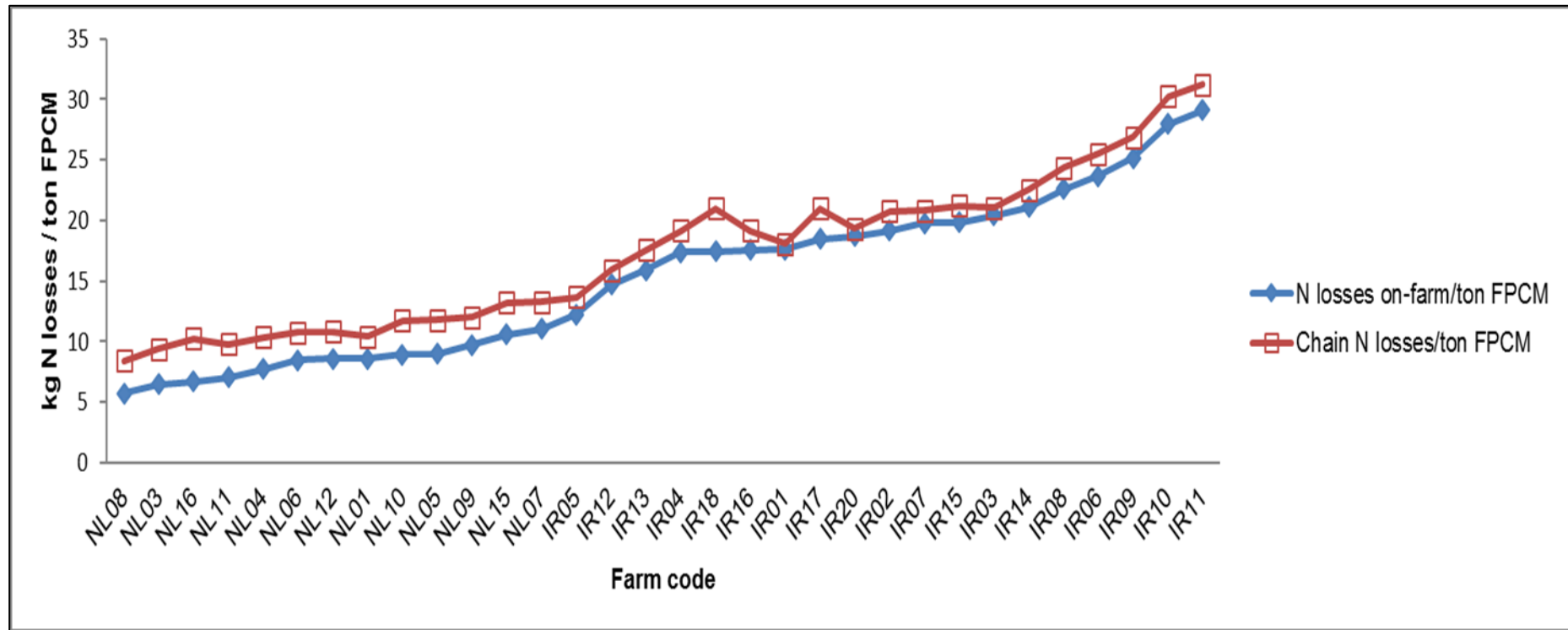|	Value	Significance (p-value)
R^2	0.99	—
β_0	0.89	0.21
β_1	2.83	0.02
β_2	1.31	0.00
β_3	-0.17	0.15

Figure 2.2 Comparison of farm rankings based on nitrogen (N) losses at farm level with N losses at chain level for 32 specialised dairy farms in the Netherlands (NL) and Ireland (IR). Farms are ranked by increasing N losses at farm level.

Phosphorus losses at farm and chain level

Table 2.6 shows mean and standard deviation of P inputs and outputs of Irish and Dutch dairy farms. Results are expressed in kg P per ha and per ton FPCM. Major P inputs for Irish farms are concentrates (7.4 kg P/ha; 0.8 kg P/ton FPCM) and mineral fertilizers (3.2 kg P/ha; 0.4 kg P/ton FPCM), whereas for Dutch farms, major P inputs are concentrates (25.5 kg P/ha; 1.1 kg P/ton FPCM) and roughages (12.4 kg P/ha; 0.5 kg P/ton FPCM). For IR farms, P inputs of organic fertilizer is 0, where for NL farms the average P inputs of organic fertilizer is negative (-13.3 kg P/ha; -0.5 kg P/ton FPCM). When expressed per ha, total inputs of the Irish farms (13.0 kg P/ha) are about half the inputs of the Dutch farms (25.4 kg P/ha). The same accounts for P outputs (i.e. 11.6 kg P/ha for IR and 23.4 kg P/ha for NL). P losses per ha, however, are similar on both farm types (3.6 kg P/ha for IR and 3.5 kg P/ha for NL). When expressed per ton FPCM, both inputs and outputs of the Irish farms (1.4 kg P/ton FPCM; 1.3 kg P/ton FPCM) are higher than the inputs and outputs of the Dutch farms (1.2 kg P/ton FPCM; 1.1 kg P/ton FPCM).

Table 2.6 Mean and standard deviation (SD) of phosphorus (P) inputs, outputs and losses of Irish and Dutch dairy farms, expresses in kg P per ha and per ton FPCM

Indicators	Per ha		Per ton FPCM	
	IR (n=19) Mean(SD)	NL (n=13) Mean(SD)	IR (n=19) Mean(SD)	NL (n=13) Mean(SD)
P inputs*				
Mineral fertilizer	3.2(3.1)	0.8(1.2)	0.4(0.4)	0.0 (0.1)
Organic fertilizer	0.0(0.0)	-13.3(11.5)	0.0(0.0)	-0.5(0.4)
Roughage	2.4(3.1)	12.4(10.0)	0.2(0.2)	0.5(0.4)
Concentrates	7.4(3.3)	25.5(10.0)	0.8(0.4)	1.1(0.2)
Total inputs	13.0(6.8)	25.4(8.6)	1.4(0.7)	1.2(0.3)
P outputs*				
Milk	8.3(1.9)	19.1(5.2)	0.9(0.0)	0.9(0.0)
Animal	3.3(1.3)	3.9(1.7)	0.4 (0.2)	0.2(0.1)
Plants	0.0(0.0)	0.4(1.1)	0.0(0.0)	0.0(0.1)
Total outputs	11.6(2.4)	23.4(6.0)	1.3(1.1)	1.1(0.1)
P losses	3.6(4.8)	3.5(4.1)	0.3(0.4)	0.1(0.2)

* All the P inputs and outputs are net inputs and outputs, which takes consideration of stock change. Organic fertilizer inputs = purchased organic fertilizer - stock change - organic fertilizer outputs. Animal outputs=animal sold-animal input + stock change. Due to lack of data on atmospheric deposition and small contribution to the total P input, in this study, we assume 0 atmospheric deposition for both IR and NL.

When expressed per ton FPCM, P losses at farm level did not differ between Irish farms (0.3 kg P/ton FPCM) and Dutch farms (0.1 kg P/ton FPCM) (Table 2.7). Although the off-farm P losses were lower on Irish farms (0.6 kg P/ton FPCM) than on Dutch farms (0.9 kg P/ton FPCM), P losses at chain level did not differ significantly between the two farm types (Table 2.7). P losses per ton FPCM are low both at farm and chain level. Some farms even have a P balance of 0 (Figure 2.3). This relates to the fact that in Ireland and the Netherlands, many soils are P-saturated, resulting in a situation where farmers rely on P that is stored in the soil during previous years. As a result, P input levels are low (i.e., changes in soil P are not accounted for), resulting in a P balance that is negative or close to zero.

Table 2.7 Mean, standard deviation (SD) and mean difference significance level (p) of phosphorus (P) losses at farm and chain level of the Irish and Dutch dairy farms, expressed per ton FPCM

Indicators[*]	Per ton FPCM		
	IR Mean (SD)	NL Mean (SD)	Mean difference (p)
P losses (kg P)			
On-farm	0.3(0.4)	0.1(0.2)	0.2(0.57)
Off-farm	0.6(0.3)	0.9(0.2)	0.3(0.00)
Chain	0.8(0.6)	1.0(0.3)	0.2(0.36)

* Indicators related to phosphorus at the chain levels are normally distributed ($P>0.05$), however, at the farm level, they are not normally distributed ($P<0.05$). Off-farm P losses per ton FPCM for both IR and NL are normally distributed.
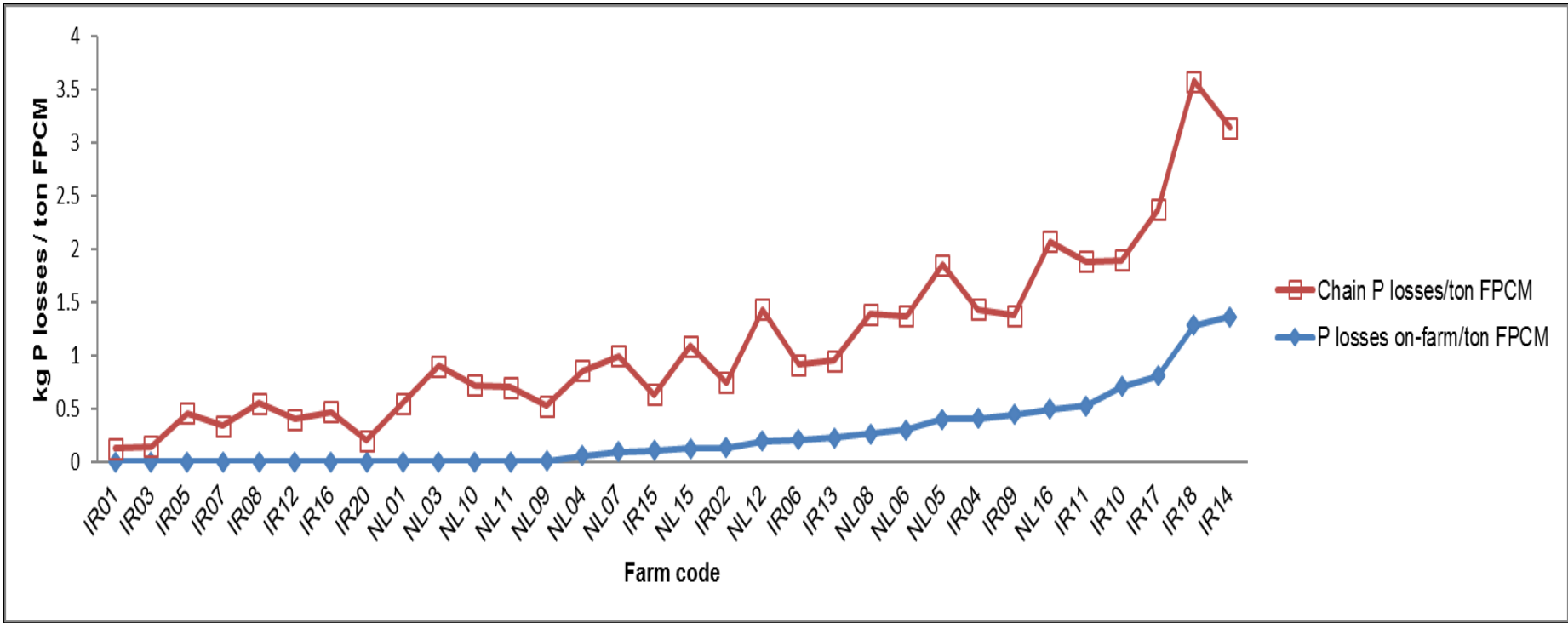
Figure 2.3 Comparison of farm rankings based on phosphorus (P) losses at farm level with P losses at chain level for 32 specialised dairy farms in the Netherlands (NL) and Ireland (IR). Farms are ranked by increasing P losses at farm level.

Table 2.8 shows results of the linear regression analysis. Off-farm P losses are higher for Dutch farms than for Irish farms (β_1=0.31, p<0.001). In general, the change in off-farm losses per unit change in on-farm losses tends to be higher on Dutch farms than on Irish farms ($\beta3$=0.49, p=0.10). As a result, the ranking pattern based on P losses at farm level differed from the ranking pattern based on P losses at chain level (Figure 2.3). The change in off-farm losses per unit differences in P losses between farm and chain level were more pronounced on Dutch farms than on Irish farms because Dutch farms rely more on purchased feed products.

Table 2.8 Results for regression of phosphorus losses at farm level on nitrogen losses at chain level.

	Value	Significance (p-value)
R^2	0.91	—
β_0	0.38	0.00
β_1	0.31	0.00
β_2	1.31	0.00
β_3	0.49	0.10

2.4 Discussion

Comparison of results with other studies

The 19 Irish dairy farms used in this study were included also in the study of Mihailescu et al. (2014), which studied 21 Irish dairy farms in total. When comparing farm area, stocking rate and milk yield per cow, the Irish farms used in this study scored close to, but slightly above the national average of dairy farms in Ireland (Mihailescu et al., 2014). The 13 Dutch farms we used in this study were included also in the study of Oenema et al. (2012), which studied 16 Dutch dairy farms in total. According to Oenema et al. (2012), these Dutch dairy farms also had, on average, a higher production intensity and milk yield per cow when compared to the national average.

The farm N balances calculated in this study (mean: IR=217.3 kg N/ha, NL=196.5 kg N/ha) are within the range reported in literature (values range from 155 kg N/ha to 295 kg N/ha) (Nevens

et al., 2006; Treacy et al., 2008; Beukes et al., 2012; Mihailescu et al., 2014). Mihailescu et al. (2014), studying dairy farms in Ireland, found an average surplus of 142 kg N/ha at farm level for 2009, a surplus of 207 kg N/ha for 2010, and a surplus of 176 kg N/ha for 2011. According to Mihailescu et al. (2014), the average N surplus per ha was higher in 2010 compared with 2009 and 2011 because in 2010 the stocking rate was higher, as well as the amount of purchased fertilizer. The latter resulted from a lower mean temperature between March and May in 2010 compared to the other two years, reducing grass growth rate and increasing the use of fertilizers. Because our objective was to examine indicators derived from a chain-based NB to benchmark contrasting dairy systems, rather than to investigate Irish farms' average performances, the high surpluses in 2010 did not influence our final conclusions. The farm P balances calculated in this study (mean: IR=3.6 kg P/ha, NL=3.5 kg P/ha) are also within the range found in literature (value range from -7 kg P/ha to 133 kg P/ha) (Spears et al., 2003; Gourley et al., 2012; Buckley et al., 2013).

The chain level NB is similar to the concept of N footprint (i.e. accounting all the losses along the chain). Leip et al. (2013), for example, calculated the N footprint of 12 main food categories and found that the N footprint of dairy products from cows, sheep and goats ranged from 10 to 45 kg N per ton of product.

Justification of assumptions

In order to compute an NB at farm level, several assumptions have to be made. The first assumption relates to the output of manure. Some studies suggest to exclude manure as an output when assessing the NUE of a farm, because manure is not an end-product for human use (Simon et al., 2000, Godinot et al., 2014). On the contrary, Spears et al. (2003) did include manure as an output when computing an NB at farm level, because manure was valued as a fertilizer. In this study, manure output was subtracted from the input of organic fertilizer. This method has recently been suggested by Godinot et al. (2014), who calculated the NUE of dairy farms at farm and chain level. When the output of manure exceeds the input of organic

fertilizer, it is considered as a negative net input instead of a positive net output. In case of calculating the NUE of a farm, manure is not valued the same as milk and meat, but it is valued for its fertilizer capacity similar to fertilizer inputs. In case of calculating the nutrient losses of a farm, however, subtracting the output of manure from the input does not affect the outcome. In other words, manure is not considered to be a loss and therefore valued the same as milk and meat. The advantage of subtracting the output of manure from the input is that losses during production of manure entering the farm, or during processing of manure leaving the farm, do not have to be considered when calculating the NB at chain level. The second assumption relates to changes in soil N and P stocks. Godinot et al. (2014) pointed out that soil N stock variation needs to be taken into account when computing an NB. They argue that excluding this variation may result in reduced soil fertility (i.e. N uptake by crops exceeds the N fertilizer input), or in overestimating N losses (i.e. part of the N surplus might be stored in the soil). In this study, however, we excluded changes in N soil stock, which implies we assumed that farms reached an equilibrium state of soil organic matter. This assumption has been used previously with the objective of simplifying the calculation processes (e.g. Powell et al., 2010; Oenema et al.,, 2012; Mihailescu et al.,, 2014). Collecting data on changes in soil organic matter content is difficult and time consuming, or requires detailed modelling techniques. Because, the objective of this paper was to investigate whether a chain level NB provides a valuable approach to benchmark systems, and methods to account for changes in soil organic matter seem unfeasible for the purpose of benchmarking due to the lack of data on soil processes across the world, we did not include changes in soil organic matter. The third assumption relates to the exclusion of N losses from energy usage. We excluded N losses related to energy usage because Godinot et al., (2014) included N losses related to the combustion of diesel in their NB and concluded that the contribution was small.

To calculate the NB at chain level, we excluded nutrient losses during production of fertilizers from our analysis, because the contribution of these losses to the total losses was found to be minor. We calculated, for instance, N losses during fertilizer production for the farm with the

highest fertilizer input, and found a loss of 0.4 kg N/ton FPCM, whereas the total loss of this farm was found the be 21.1kg N/ton FPCM. Including this loss, therefore, will not significantly affect the overall NB of the farm.

Nutrient losses at chain level were expressed per ton FPCM, while nutrient losses at farm level were expressed per ton FPCM and per ha of land. Nutrient losses per unit of product are a measure for the efficiency of production, whereas nutrient losses per ha of land are a measure for the potential local impact of a farm (or another system). Nutrient losses at chain level, therefore, cannot be expressed per ha, because purchased feed ingredients are produced at different locations and different countries. Computing one average value for nutrient losses 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. In addition, without site specific knowledge on the local ecosystems that are exposed to the pollutants, it's not possible to quantify the local impacts accurately (Potting and Hauschild, 2006). We suggest, therefore, to use indicators at different hierarchical levels: a product-based indicator, such as an N loss along the chain which provides insight into the efficiency of production, and a region-based indicator, such as an N loss per ha at regional level, which provides insight into the local impact of production.

Interpretation of nutrient balance

Gerber et al. (2014) pointed out that the limitation of an NB approach is that the farm or system is considered as a black box and internal flows are not analysed. Despite this limitation, however, an NB can still be a valuable approach to benchmark farms rather than to diagnose detailed problems existing in the farming practice (Powell et al., 2010; Gourley, et al., 2012; Toma et al., 2013). Although a positive NB at farm level does not imply actual losses, it still provides insights in the potential losses in the future (Halberg, 1999, cited in Steinshamn et al., 2004). A positive N balance was defined as an N loss, potentially contributing to environmental pollution through e.g., the emission of nitrous oxide, the volatilisation of

ammonia and nitrogen oxide, and through runoff and leaching of dissolved nitrate (Gourley et al., 2012). A negative N balance means that the soil is depleted. Altieri and Pengue (2006), for example, indicated that intensive soybean cultivation in Latin America resulted in severe depletion of soil nutrients and reduced soil fertility. Our results also show a negative N balance for soybeans produced in Argentina (i.e. 33% of the soybeans in our study were assumed to come from Argentina). Because a negative N balance is not as a loss, we assumed the loss is 0 in this situation.

Interpretation of a P balance can be difficult, because P can be stored and immobilised in the soil, and can remain stored in the soil over a long time. This means that to a certain extent, P is variously available to plants independent on the year of application (Buckley, et al., 2013). A positive P balance does not necessarily mean an immediate loss to the environment, because soils with a low P-status can build P-reserves up to an optimum level (Bomans et al., 2005). Only in the case that a P surplus leads to soil P levels above agronomic levels, the P surplus can pose a risk of P losses in surface water (Sharpley, 1995, as cited in Gourley et al., 2012). The same principle can be applied for negative P balances. Several Irish and Dutch farms show a negative P balance at farm level. This negative P balance was not considered to be a loss, but set to zero. In Ireland and the Netherlands, soils are commonly P-saturated. As a result, farmers reduce or stop applying mineral P fertilizer on the field. If we only look at the farm balance on a yearly basis, the P balance of such a farm can be negative, but it does not imply a P deficit (Humphreys et al., 2009; Buckley et al., 2012). However, negative P balance may not be a problem on short-term, but it can cause problems if soil P is already below adequate agronomic levels on the long-term (Steinshamn et al., 2004; Gourley et al., 2012). Soil-tests that can monitor P levels need to be performed regularly to maintain an adequate P status of the soil.

The value of a nutrient balance at chain level

When comparing dairy production systems, a farm level NB might lead to biased results, because differences in nutrient losses related to off-farm production processes are not taking into account. We compared N losses per ton FPCM of Irish and Dutch farms at farm and chain level, and demonstrated that Dutch farms are more efficient than Irish farms at both levels. The differences in on-farm N losses per ton FPCM between Irish and Dutch farms are so big, that the relatively small differences in off-farm nutrient losses do not alter the conclusions drawn at farm level (Table 2.4). Furthermore, because the change in off-farm losses per unit change in on-farm losses was similar for both farm types, ranking of farms based on N losses at farm level was similar to ranking based on N losses at chain level. We conclude, therefore, that in our dataset, N losses at farm level can be used to benchmark N efficiency of farms. When comparing P losses per ton FPCM of Irish and Dutch farms, we found that Irish farms are not different from Dutch farms at both levels. Ranking of farms based on P losses at farm level, however, appeared different from ranking based on P losses at chain level. This difference in ranking resulted from the relatively high importance of off-farm losses compared to on-farm losses, and from the fact that the change in off-farm losses per unit change in on-farm losses differed between farms. Our results show that in general, farms (e.g. farms with a high input of, e.g., feed) that are relatively efficient at farm level could be less efficient at chain level, and vice versa. To benchmark P efficiency of farms, therefore, an NB at chain level is required as an NB at farm level could lead to biased conclusions.

2.5 Conclusion

Although the importance of computing nutrient losses at the chain level has been addressed before, to the best knowledge of the authors, this study is the first to use farm data to quantify and compare nutrient losses at farm and chain level, in order to determine the relevance of a chain approach to benchmark milk production systems. We concluded that to benchmark systems, a nutrient balance at farm level can be used if differences in on-farm losses between systems are large, and off-farm losses are relatively unimportant. To benchmark individual farms, a nutrient balance at farm level can be used if changes in off-farm losses per unit change

in on-farm losses are similar across farms. A chain level balance of a sample set, however, is required to verify these conditions. Results of this study can be used to inform decision makers about the level of analysis that is required to benchmark nutrient losses of farming systems.

2.6 References

Altieri, M. , Pengue, W., 2006. GM soybean: Latin America's new coloniser. Seedling. 13–17.

Buckley, C., Murphy, P. and 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.

Bomans, E., Fransen, K., Gobin, A., Mertens, J., Michiels, P., Vandendriessche, H., Vogels, N., 2005. Addressing phosphorus related problems in farm parctice. Final report to the European Commission. Soil service of Belguim.

Boland, F., Grady, L. O. , More, S. J. , 2013. Investigating a dilution effect between somatic cell count and milk yield and estimating milk production losses in Irish dairy cattle. J. Dairy. Sci. 96(3), 1477–1484.

Bengtsson, H., O'born, I., Jonsson, S., Nilsson, S.I., Andersson, A., 2003. Field balances of some mineral nutrients and trace elements in organic and conventional dairy farming—a case study at O¨ jebyn, Sweden. Eur. J. Agron. 20, 101-116.

Beukes P.C., Scarsbrook M.R., Gregorini P., Romera A.J., Clark D.A., Catto W., 2012. The relationship between milk production and farm-gate nitrogen surplus for the Waikato region, New Zealand. J. Envrion. Manage. 93, 44–51.

CVB, 2008. Chemische samenstelling en nutritionele waarden van voedermiddelen. Productschap Diervoeder, Den Haag, the Netherlands. ISSN1567-8679.

Dairyman,2010. < http://www.interregdairyman.eu/en/dairyman.htm>

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.

Djekic, I., Miocinovic, J., Yomasevic, I., Smigic, N., Tomic N., 2014. Environmental life-cycle assessment of various dairy products. J. Clean. Prod. 68, 64–72.

Dolman, M.A., Sonneveld, M.P.W., Mollenhorst, H., de Boer, I.J.M.,2014. Benchmarking the economic, environmental and societal performance of Dutch dairy farms aiming at internal recycling of nutrients J. Clean. Prod. 73, 245–252.

EMEP, 2007. Co-operative programme of the monitoring and evaluation of the long range transmissions of air pollutants in Europe. <http://www.nilu.no/ projects/ccc/emepdata.html>

EU, 2010. The European commission. Official journal of the European Union, 6.2.2010. Commission Decision, amending Decision 2005/88o/EC granting a derogation requested by the Netherlands pursuant to Council Directive 91/676/EEC concerning the protection of waters against pollution caused by nitrates from agricultural sources (2010/65/EU). Notified under document C (2010).606

European Communities, 2010. (Good Agricultural practice for Protection of Waters) Regulations 2010, S.I. 610 of 2010. Dublin, Stationery Office, Ireland.

Food and Agriculture Organisation (FAO), 2006. FAO World agriculture: Towards 2030/2050 Interim Report Food and Agriculture Organisation, Rome.

Gerber, P.J., Uwizeye, A., Schulte, R.P.O., Opio, C.I., de Boer, I.J.M., 2014. Nutrient use efficiency: a valuable approach to benchmark the sustainability of nutrient use in global livestock production? Curr Opin Environ Sustain 9–10,122–130.

Gourley, C.J.P., Powell, J.M., Dougherty, W.J., Weaver, D.M., 2007. Nutrient budgeting as an approach for improving nutrient management on Australian dairy farms. Aust. J. Exp. Agric. 47, 1064–1074.

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.I., 2012. Farm-scale nitrogen, phosphorus, potassium and sulphur balances and use efficiencies on Australian dairy farms. Anim. Prod. Sci. 52, 929–944.

Godinot, O., Carof, M., Vertes, F., Leterme, P., 2014. SyNE: An improved indicator to assess nitrogen efficiency of farming systems. Agric. Syst. 127, 41-52.

Garnett, T., Appleby, M.C., Balmford, A., Bateman, I.J., Benton, T.G., Bloomer, P.,Burlingame, B., Dawkins, M., Dolan, L., Fraser, D., Herrero, M., Hoffmann, I.,Smith, P., Thornton, P.K., Toulmin, C., Vermeulen, S.J., Godfray, H.C.J., 2013. Sustainable intensification in agriculture: premises and policies. Science. 341, 33–34.

Humphreys, J., Aarts, H.F.M., Watson, C.J., Wachendorf, M., Le Gall, A., Taube, F., Pfimlin, A., 2009. Sustainable options for grassland-based dairy production in the Northwest of Europe. Ir. J. Agri-Environ. Res. 7, 1-14.

Halberg, N., 1999. Indicators of resource use and environmental impact for use in a decision aid for Danish livestock farmers. Agr. Ecosyst. Environ, 76, 17–30.

Halberg, N., van der Werf, H., Basset-Mens, C., Dalgaard, R., de Boer, I.J.M., 2005. Environmental assessment tools for the evaluation and improvement of European livestock production systems. Livest. Prod. Sci. 96, 33–50.

Huhtanen, P., Nousiainen, J., Turtolad. E., 2011. Dairy farm nutrient management model: 2. Evaluation of different strategies to mitigate phosphorus surplus. Agric. Syst.104, 383–391.

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.

Mihailescu, E., Murphy, P.N.C., Ryan, W., Casey I. A., Humphreys. J., 2014. Nitrogen balance and use efficiency on twenty-one intensive grass-based dairy farms in the South of Ireland. J. Agr. Sci. doi:10.1017/S0021859614000045.

Mont, O., Heiskanen, E., Neuvonen, A., Lähteenoja, S., 2014. Sustainable lifestyles 2050: stakeholder visions, emerging practices and future research. J. Clean. Prod. 63, 24–32.

Nevens, F., Verbruggen, I., Reheul, D., Hofman. G., 2006. Farm gate nitrogen surpluses and nitrogen use efficiency of specialized dairy farms in Flanders: Evolution and future goals. Agric. Syst. 88, 142–155.

Oenema, O., Kros, H., deVries, W., 2003. Approaches and uncertainties in nutrient budgets: implications for nutrient management and environmental policies. Eur. J. Agric. 20, 3–16.

Oenema, J., Koskamp, G.J., Galama, P.J., 2001. Guiding commercial pilot farms to bridge the gap between experimental and commercial dairy farms: the project 'Cows & Opportunities'. Neth. J. Agr. Sci. 49, 277-296.

Oenema, J., van Ittersum, M., van Keulen, H., 2012. Improving nitrogen management on grassland on commercial pilot dairy farms in the Netherlands. Agr. Ecosyst. Environ. 162, 116–126.

Product Board Animal Feed (Productschap Diervoeder).2008.Tabellenboek Veevoeding 2008. Productschap Diervoeder, CVB, Den Haag.

Powell, J.M., Gourley, C.J.P., Rotz, C.A., Weaver, D.M., 2010. Nitrogen use efficiency: a measurable performance indicator for dairy farms. Environ. Sci. Policy. 13, 217–228.

Potting, J., Hauschild, M.Z., 2006. Spatial differentiation in life cycle impact assessment: a decade of method development to increase the environmental realism of LCIA. Int. J. Life. Cycle. Assess. 11, 11–13.

Raison, C., Pflimlin, A., le Gall, A., 2006. Study of the distribution of nitrogen surplus in experimental dairy farms of the Atlantic area. Dairy Systems and Environment in the Atlantic Area. Proceedings of the Final Seminar of the Green Dairy Project, Institute of de l'Elevage, Rennes, France, 67–95.
SPSS, 2011. Statistical Package for Social Sciences. SPSS Software Package 20.0 for Microsoft Windows, Chicago, Illinois, USA.

Sharpley, A.N., 1995. Dependency of runoff phosphorus on extractable soil phosphorus. J. Environ. Qual., 24, 920–926.

Sterk A, personal communication 2014.

Steinshamn, H., Thuen, E., Azzaroli Bleken, M., Tutein Brenøe, U., Ekerholt, G., Yri, C., 2004. Utilization of nitrogen (N) and phosphorus (P) in an organic dairy farming system in Norway. Agr. Ecosyst. Environ. 104 , 509–522.

Simon, J.C., Grignani, C., Jacquet, A., Corre, L.L., Pagès, J., 2000. Typologie des bilans d'azote de divers types d'exploitation agricole: recherche d'indicateurs de fonctionnement. Agronomie.20, 21.

Sutton, J.N., Johannessen, S.C., Macdonald, R.W., 2013. A nitrogen budget for theStrait of Georgia, British Columbia. Biogeosci. Discuss. 10, 7135–7169.

Spears, R.A., Kohn, R.A., Young, A.J., 2003. Whole-farm phosphorus balance on western dairy farms. J. Dairy. Sci. 86, 688–695.

Toma, L., March, M., Stott, A.W., Roberts, D.J., 2013. Environmental efficiency of alternative dairy systems: A productive efficiency approach. J. Dairy. Sci. 96, 7014-4031.

Thomassen, M.A., de Boer, IJ.M., 2005. Evaluation of indicators to assess the environmental impact of dairy production systems. Agr. Ecosyst. Environ. 111,185–199.

Treacy, M., Humphreys, J., McNamara, K., Browne, R., Watson, C.J., 2008. Farm-gate nitrogen balances on intensive dairy farms in the south west of Ireland. Irish. J. Agr. Food. Res. 47, 105–117.

UN News Centre., 2013. World population projected to reach 9.6 billion by 2050 – UN report.[WWW Document]. < http://www.un.org/apps/news/story.asp?NewsID=45165> (accessed11.09.2014).

Vellinga, T.V., Blonk, H., Marinussen, M., van Zeist, W.J., de Boer, I.J.M., Starmanset, D., 2013. Methodology used in FeedPrint: a tool quantifying greenhouse gas emissions of feed production and utilization. Report 674. Wageningen UR Livestock Research. <http://edepot.wur.nl/254098>

Volk, M., Liersch, S., Schmidt, G., 2009. Towards the implementation of the European Water Framework Directive? Lessons learned from water quality simulations in an agricultural watershed. Land Use Policy, 26, 580–588.

Chapter 3

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

W. Mu[1,2*], C.E. van Middelaar[1], J.M. Bloemhof[2] , B. Engel[3], I.J.M. de Boer[1]

[1] Animal Production Systems group, Wageningen University, P.O. Box 338, 6700 AH Wageningen, the Netherlands

[2] Operations Research and Logistics group, Wageningen University, P.O. Box 8130, 6700 EW Wageningen, the Netherlands

[3] Biometris, Wageningen University, P.O. Box 16, 6700 AA Wageningen, The Netherlands

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.

3.1 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 80% 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 to 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 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, C. 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 (Lebacq 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.

3.2 Material and methods

3.2.1 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 3.1).

Table 3.1 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

3.2.2 System boundaries

Figure 3.1 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).
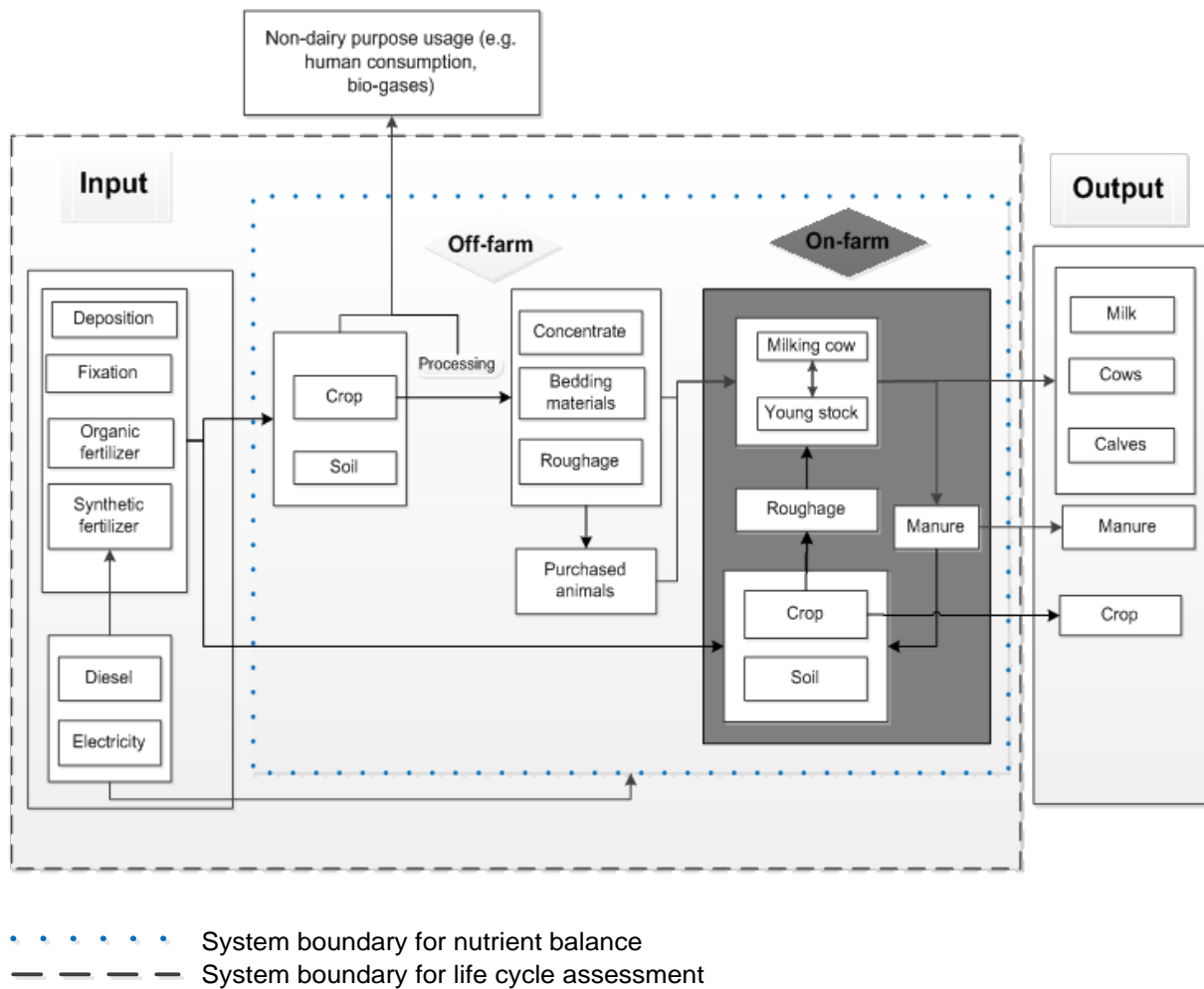
Figure 3.1 System boundaries

3.2.3 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 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) x [0.337 + 0.116 x Fat (%) + 0.06 x Protein (%)] (Product Board Animal Feed 2008)). Nutrient surpluses were allocated to milk and meat based on the relative economic value (i.e. economic allocation).

3.2.4 Life cycle assessment

This LCA study focussed on the use of scarce recourses land and energy, and on the emission of pollutants methane (CH_4), nitrous oxide (N_2O), carbon dioxide (CO_2), ammonia (NH_3), mono-nitrogen oxides (NO_x), sulphur dioxide (SO_2), nitrate (NO_3^-) and phosphate (PO_4^{3-}). The impact categories considered were land use, fossil energy use, climate change, acidification, freshwater eutrophication, and marine eutrophication (Table 3.2).

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 litre of milk), electricity use for crop production (37 kWh per ha), and diesel use for crop production (128 litre per ha). In addition, diesel use for forage and pasture production was dependent on the area of maize silage (77 litre per ha for areas with < 5% maize silage; 94 litre per ha for areas with 5-25% maize silage; and 128 litre 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, CH_4 emissions were based on IPCC (2006) Tier 2. Enteric CH_4 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 CH_4 from manure management were based on the 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). CH_4 per kg volatile solid excreta was calculated by multiplying the maximum amount of 0.36 kg CH_4/kg (dairy cow) or 0.12 kg CH_4/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 N_2O emissions occur via direct and indirect pathways. Direct N_2O emissions result from manure management and N application to the field, whereas indirect N_2O emissions results from volatilization of NH_3, NO_x, and leaching of NO_3^- (IPCC, 2006). Direct N_2O emissions from manure storage were based on the N excretion multiplied by a management specific emission factor (i.e., 0.002 kg N_2O-N per kg N for manure in pit storage, and 0.005 kg N_2O-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 N_2O emissions from N application via manure, synthetic fertilizer, and crop residues, and all indirect N_2O emissions were based on IPCC Tier 1 (IPCC, 2006).

On-farm NH_3 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 NH_3-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 NH_3 related to the application of synthetic fertilizer were based on fertilizer specific emission factors (EEA, 2013).

On-farm emissions of NO_x and SO_2 related to the combustion of diesel were based on Eco-invent (2010).

Due to lack of detailed soil information, we estimated leaching of NO_3^- and PO_4^{3-} 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 NO_3^- was calculated by subtracting NH_3-N, N_2O-N and NO_x-N from the total N surplus of the farm, whereas for PO_4^{3-}, 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.

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 CO_2, CH_4 and N_2O. Acidification of terrestrial and aquatic ecosystems relates to the emission of

acidifying gases, including SO_2, NH_3 and NO_x (i.e., NO and NO_2). Eutrophication of terrestrial and aquatic ecosystems relates to leaching of eutrophic substrates, including NO_3^- and PO_4^{3-}, and to the emission of eutrophic gases, including NH_3 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 3.2. 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.

Table 3.2 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_2O	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	

3.2.5 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 indicators. 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, 1998). 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).

3.3 Results

3.3.1 Nutrient balance

Indicators derived from the NB are presented in Table 3.3. 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).

Table 3.3 Indicators derived from the nutrient balance

Indicator	Unit	Mean	SD	Minimum	Maximum
N surplus					
On-farm	g N x kg FPCM^{-1}	12.2	5.3	3.5	25.0
Off-farm	g N x kg FPCM^{-1}	2.3	1.3	0.4	6.2
Total	g N x kg FPCM^{-1}	14.4	5.2	4.9	27.0
P surplus					
On-farm	g P x kg FPCM^{-1}	0.5	0.7	0.0	3.0
Off-farm	g P x kg FPCM^{-1}	0.8	0.4	0.1	1.9
Total	g P x kg FPCM^{-1}	1.3	0.9	0.1	4.9

3.3.2 Life cycle assessment

Indicators derived from the LCA are presented in Table 3.4. Producing one kg FPCM requires on average 1.3 m^2 land and 2.8 MJ energy, and results in a GWP of 1.2 kg CO$_2$-eq, an AP of

26.1 g SO_2-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 activates, 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 3.4 Indicators derived from the LCA

Indicator	Unit	Mean	SD	Minimum	Maximum
Land use	m^2 x kg FPCM^{-1}				
On-farm		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
Energy use	MJ x kg FPCM^{-1}				
On-farm		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
GWP[1]	kg CO_2 eq x kg FPCM^{-1}				
On-farm		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
AP[1]	g SO_2 eq x kg FPCM^{-1}				
On-farm		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
FEP[1]	g P eq x kg FPCM^{-1}				
On-farm		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
MEP[1]	g N eq x kg FPCM^{-1}				
On-farm		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

[1.] GWP (global warming potential), AP (acidification potential), FEP (freshwater eutrophication potential), MEP (marine eutrophication potential)

3.3.3 Statistical analysis

Correlations are presented in Table 3.5. 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).

Table 3.5 Pearson correlation coefficient matrix of the environmental indicators [***]

	N surplus	P surplus	Land use	Energy use	GWP[1]	AP[1]	FEP[1]	MEP[1]
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[1]					-	0.81[**]	0.28[*]	0.51[**]
AP[1]						-	0.19	0.42[**]
FEP[1]							-	0.41[**]
MEP[1]								-

[1.] GWP (global warming potential), AP (acidification potential), FEP (freshwater eutrophication potential), MEP (marine eutrophication potential)

[*] 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.

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 3.6. 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 3.6).

Table 3.6 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[1]	0.95
GWP[1]	0.84
Energy use	0.82
Canonical correlation	
N surplus, FEP[1], GWP[1], energy use	0.98

[1] FEP (freshwater eutrophication potential), GWP (global warming potential)

[*] All correlation coefficients are significant at the 0.05 level (2-tailed)

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 3.7. 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 3.6). 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.

Table 3.7 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[1]	0.95
MEP[1]	0.91
GWP[1]	0.83
AP[1]	0.79
Canonical correlation	
FEP, MEP, GWP, AP	0.97

[1] FEP (freshwater eutrophication potential), MEP (marine eutrophication potential), GWP (global warming potential), AP (acidification potential)

[*] All correlation coefficients are significant at the 0.05 level (2-taiiled)

3.4 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; Lebacq 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, measureable, 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 indicators. 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.

3.5 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.

Acknowledgement

3.6 Reference

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.

de Boer, I.J.M., 2003. Environmental impact assessment of conventional and organic milk production. Livest Prod Sci. 80, 69–77.

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., 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, Goteborg, Sweden.

Cederberg, C., Flysjö, A., 2004. Life Cycle Inventory of 23 Dairy Farms in South-Western SweSIK-report 728 Swedish Institute for Food and Biotechnology, SIK, Gothenburg, Sweden.

Cohen, J., 1988. Statistical Power Analysis for the Behavioral Sciences, 2nd edn. Lawrence Erlbaum Associates Inc, Abingdon.

Dairyman,2010. < http://www.interregdairyman.eu/en/dairyman.htm>

Eco-invent, 2010. Eco-invent data v2.2. Final reports eco-invent. Swiss Centre for Life Cycle Inventories, Dübendorf

European Environment Agency (EEA), 2013. EMEP/EEA Air Pollutant Emission Inventory Guidebook. http://www.eea.europa.eu/publications/emep-eea-guidebook-2013 [accessed 08-03-2016]

Eurostat, 2015. http://ec.europa.eu/eurostat/statistics-explained/index.php/Air_pollution_statistics [accessed 08-03-2016]

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-melk-or-eggs.htm [accessed 03-08-2016]

Gac, A., Deltour, L., Cariolle, M., Dolle, J.B., Espagnol, S., Flenet, F., Guingand, N., Lagadec, S., Le Gall, A., Lellahi, A., Malaval, C., Ponchant, P., Taileur, 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., Falcucci, 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.I., 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. Operational guide to the ISO standards. Kluwer Academic Publishers, Dordrecht.

IPCC, 2006. Intergovernmental Panel on Climate Change. Guidelines for national greenhouse gas inventories. Volume 4: agriculture, forestry and other land use. Prepared by the National Greenhouse Gas Inventories Program. Eggleston HS, Buendia L, Miwa K, Ngara T, Tanabe K (eds). 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.

Lebacq, 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.

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.

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., H. Kros, and W. de Vries. 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

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 08-03-2016]

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.

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.

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.

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.

Chapter 4

Benchmarking nutrient use efficiency of dairy farms: the effect of epistemic uncertainty

W. Mu[1*], E.A. Groen[1*], C.E. van Middelaar[1], E.A.M. Bokkers[1], S. Hennart[2], D. Stilmant[2], I.J.M. de Boer[1]

[1] Animal Production Systems group, Wageningen University, P.O. Box 338, 6700 AH Wageningen, the Netherlands

[2] Farming Systems, Territory and Information Technologies Unit, Walloon Agricultural Research Centre, 100 rue du Serpont, 6800 Libramont, Belgium

[*] Combined first authorship

Abstract

The nutrient use efficiency (NUE) of a system, generally computed as the amount of nutrients in valuable outputs over the amount of nutrients in all inputs, is commonly used to benchmark the environmental performance of dairy farms. Benchmarking the NUE of farms, however, may lead to biased conclusions because of differences in major decisive characteristics between farms, such as soil type and production intensity, and because of epistemic uncertainty of input parameters caused by errors in measurement devices or observations. This study aimed to benchmark the nitrogen use efficiency (NUE_N; calculated as N output per unit of N input) of farm clusters with similar characteristics while including epistemic uncertainty, using Monte Carlo simulation. Subsequently, the uncertainty of the parameters explaining most of the output variance was reduced to examine if this would improve benchmarking results. Farms in cluster 1 (n=15) were located on sandy soils and farms in cluster 2 (n=17) on loamy soils. Cluster 1 farms were more intensive in terms of milk production per hectare and per cow, had less grazing hours, and fed more concentrates compared to farms in cluster 2. The mean NUE_N of farm in cluster 1 was 43%, while in cluster 2 it was 26%. Input parameters that explained most of the output variance differed between clusters. For cluster 1, input of feed and output of roughage were most important, whereas for cluster 2, the input of mineral fertilizer (or fixation) was most important. For both clusters, the output of milk was relatively important. Including the epistemic uncertainty of input parameters showed that only 37% of the farms in cluster 1 (out of 105 mutual comparisons) differed significantly in terms of their NUE_N, whereas in cluster 2 this was 82% (out of 120 comparisons). Therefore, benchmarking NUE_N of farms in cluster 1 was no longer possible, whereas farms in cluster 2 could still be ranked when uncertainty was included. After reducing the uncertainties of the most important parameters, 72% of the farms in cluster 1 differed significantly in terms of their NUE_N, and in cluster 2 this was 87%. Results indicate that reducing epistemic uncertainty of input parameters can significantly improve benchmarking results. The method presented in this study, therefore, can be used to draw more reliable conclusions regarding benchmarking the NUE of farms, and to identify the parameters that require more precision to do so.

4.1 Introduction

Nitrogen (N) is an essential nutrient for milk production. The input of N into European milk production systems has increased in the past decades, mainly via purchase of fertilizer and feed, but also via atmospheric deposition and biological fixation (Powell et al., 2010). These increased N inputs have also increased N losses to the environment, via leaching of nitrate (NO_3^-) and emissions of N-gases, such as nitrous oxide (N_2O) and ammonia (NH_3). These N losses contribute to environmental problems, such as eutrophication, acidification and global warming (Whitehead., 1995; Smith et al., 1999). To tackle this problem, the European Union introduced legislation, such as the Nitrates Directive (EU, 2006), which set limits on N application per hectare to reduce NO_3^- leaching.

There has been on-going studies and discussions on how to reduce N losses of dairy farms in Europe (e.g. Aarts et al., 1992, Schröder et al., 2003, Nevens et al., 2006, Phuong et al., 2013; Mihailescu et al., 2015). Calculating the nutrient balance at farm level is the most commonly used approach to evaluate these losses. In the Netherlands, for example, dairy farms are obliged to quantify their annual nitrogen and phosphorus balance from 2016 onwards (Veeteelt, 2015). A nutrient balance reflects the difference in nutrients entering and leaving a system, and allows computation of environmental indicators, such as the nutrient use efficiency (NUE) or the nutrient surplus per ha of a farming system (Spears et al., 2003). NUE generally is computed as the amount of nutrients in valuable outputs of a system over the amount of nutrients in all inputs of that system (Nevens et al., 2006).

Due to the simplicity of the method and relatively low data requirement, the nutrient balance has been used as a tool to benchmark the environmental performance of farms (Oenema et al., 2003, Schröder et al., 2003). Benchmarking is defined by Camp et al (1989) as "the search for those best practices that will lead to the superior performance" and, in this study, relates to the comparison of farms based on their environmental performance in order to identify differences and potentially, improvement options. Benchmarking farms based on, for example,

their NUE, however, may lead to biased conclusions because of two reasons. First, as pointed out by Schröder et al. (2003), comparing the NUE of farms is justified only if they have similar major decisive characteristics. These characteristics can be based on: (unmanageable) physical factors, such as soil type and climatic conditions (Roberts, 2008; Powell, et al., 2010); and long term strategic decisions, such as the degree of self-sufficiency (e.g. grass-based versus concentrate-based), production intensity, or manure management system (Nevens et al., 2006). Other characteristics that have an influence of the NUE of a farm include short term tactical decisions, such as choice of the feed crop, or grazing regime; operational decisions (i.e., day to day decisions); and other management skills of the farmer, such as the capacity to reduce losses (e.g. losses of feed, nutrients, milk or cows (culling)) (Nevens et al., 2006). Benchmarking NUE of farms should be based on differences in short term strategic and tactical decision-making, rather than differences in physical factors and long term decisions. Second, comparing NUE of farms may be affected by epistemic uncertainty of input data, caused by errors in measurement devices or errors around observations. Epistemic uncertainty can arise from e.g. errors in practically determining the N fixation by clover, measurement errors around the feed intake of the cows or estimations around the N-content of the animals (Oenema et al., 2015). Increasing knowledge or better measurements can reduce epistemic uncertainty (Walker et al., 2003; Groen, 2016).

Previous studies focused on examining the epistemic uncertainties of nutrient flows by looking into e.g. quantity of nutrient inputs (Mulier et al., 2003; Gourley et al., 2012; Oenema et al., 2015). However, they did not examine the impact of epistemic uncertainties on benchmarking results, nor did they benchmark farms with similar decisive farm characteristics.

The objectives of this study were to benchmark the nutrient losses by comparing nitrogen use efficiency (NUE_N) of farms with similar decisive characteristics while including epistemic uncertainty, and to examine which input parameters explain most uncertainty of NUE_N results. In addition, the epistemic uncertainties of input parameters that explain most of the output variance were reduced, to illustrate how this will improve benchmarking results.

4.2 Materials and methods

4.2.1 Case study: European specialized dairy farms

We used data of specialized dairy farms from Dairyman. Dairyman was a project directed at improving regional prosperity through better resource utilization on 113 dairy farms in different European countries (Dairyman, 2010). From the 113 farms, 32 specialized dairy farms were selected. Specialized dairy farms were defined 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. These 32 dairy farms were located in different countries and regions (i.e. Netherlands (7), Ireland (13), Belgium (Flanders 8, Wallonia 2), Germany (1) and Luxembourg(1)). Selected dairy farms differed in soil types (i.e. sandy soil, loam soil), milk production (i.e. milk production per cow and per ha), grazing hours per year, and feed import (i.e. kg concentrate usage per cow per year; Table 4.1). Whereas data on soil type, milk production and feed import were based on measured farm data, data on grazing hours per year were based on estimations by the farmers. Farm data from the year 2010 were used as baseline values to determine all N-flows.

Table 4.1 Characteristics of the 32 European specialized dairy farms used in this study.

Characteristics	Unit	Mean	Minimum	Maximum
Agricultural area	ha	65	25	270
Herd size	number of dairy cows	90	37	384
Milk production	kg milk cow^{-1} $year^{-1}$	7689	5700	9853
Milk production	kg milk ha^{-1} $year^{-1}$	12598	3448	26300
Grazing hours	hour $year^{-1}$	2857	0	5146
Concentrate usage	kg cow^{-1} $year^{-1}$	1215	317	2459

4.2.2 Defining homogenous farm clusters

To enable benchmarking of NUE_N of farms with similar characteristics, farms were sorted into homogenous groups (i.e. typologies) based on their characteristics (Table 4.1). For this

purpose, we used a two-step cluster analysis, because it allows using both continuous and categorical variables as clustering criteria (Chiu et al., 2001). To perform a cluster analysis with n criteria, a sample size of 2^n farms is required (Formann, 1984). Since our sample size included 32 farms, we selected 5 criteria for the cluster analysis, namely grazing hours, soil type, concentrate per cow per year, milk production per cow per year and milk production per ha (De Vries et al., 2015, Daatselaar et al., 2015). The analysis was performed in the statistical software package IBM SPSS statistics 22 (SPSS, 2015).

4.2.3 System boundary and model assumptions of calculating NUE_N

The NUE_N was quantified at farm level, implying that only on-farm flows and losses were considered. The N-flows through a dairy farm included in this study are visualized in Figure 4.1. Inputs of N include N in mineral fertilizers, manure, animals, concentrates, roughages, biological N fixation and atmospheric N deposition. Outputs of N include N in animals, milk, manure and roughage. Stock changes (defined as final stock minus initial stock) of the mineral fertilizers, manure, animals, concentrates and roughages were taken into consideration during the computation processes. Manure output was subtracted from the total fertilizer input (i.e. through mineral fertilizer and manure). If the total manure output of the farm exceeded its total fertilizer input, excessive manure was treated as a loss. The internal N-flow from crop production to feed storage was based on the energy requirements of the herd, minus feed input and stock changes of feed. The calculation rules are specified in the supplementary material. Losses of N from manure storage were based on storage type (i.e. slurry, solid) and the baseline values of manure N in all calculations(EEA, 2013).
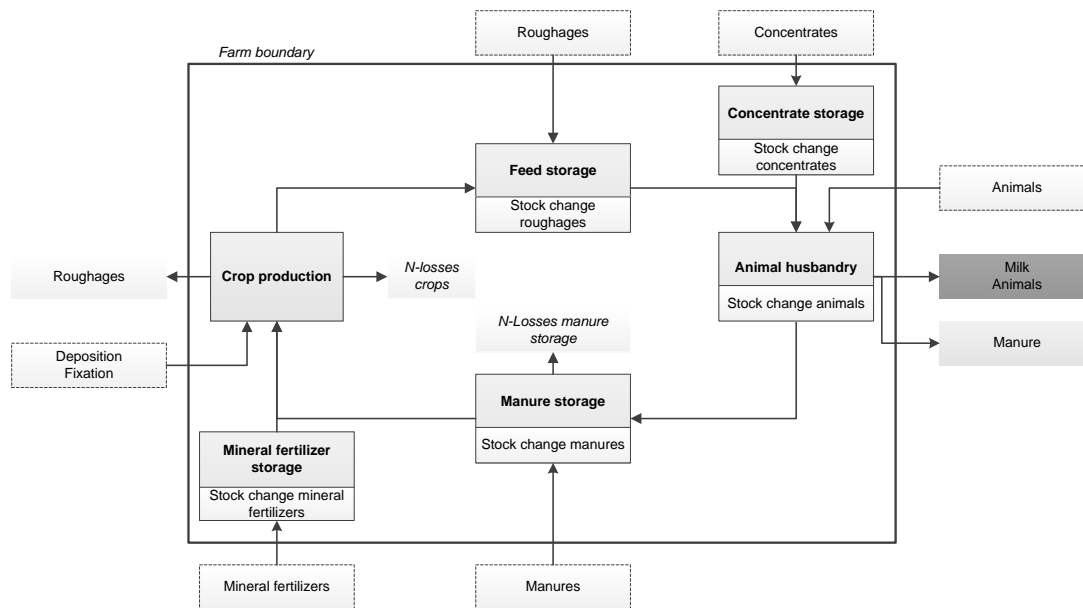
Figure 4.1 N-flows on a dairy farm to assess nutrient use efficiency; the production processes are given by the solid boxes, the N-flows are given by the arrows. A detailed description of the input parameters can be found in Table 4.2.

4.2.4 Matrix based calculation for on-farm NUE_N

We used the matrix-based approach developed by Suh and Yee (2011) to quantify the N-efficiency of the 32 dairy farms. This approach was used to describe the herd and crop balance (Figure 4.1) in one equation, which facilitates the global sensitivity analysis to examine epistemic uncertainty. A matrix-based approach allows for the presence of loops and parallel components, as is often the case on dairy farms (e.g. manure is used for the production of feed crops, which are consequently fed to the animals, producing manure). This approach requires a detailed insight into the nutrient flows within the farm.

The difference between the matrix-based approach to assess the farm N-balance and the common nutrient balance approach is that in case of the matrix-based approach the internal flows are considered (e.g. the flows between manure storage and crop production, or crop production and feed storage), just as in a substance flow analysis. In the common nutrient balance, the farm is considered as a black box (e.g. Oenema et al.,2015 and Mu et al., 2015). For more details, see the supplementary material.

In the matrix-based approach, the internal N-flows in Figure 4.1 are described by the **V** and **U**-matrix, where the **V**-matrix describes how much kg N is supplied to each production process. The **U**-matrix describes how much kg N is used by each production process (Suh and Yee, 2011). The N-flows are corrected for the stock changes (**s**) on the farms. Combined, they are quantified in a matrix **A** for each (intermediate) process. T refers to the transpose. The vector (**b**) gives the amount of nutrients extracted (**r**) to produce 1 unit of final product, which, in this case, is determined by the valuable outputs of the farm:

$$\mathbf{b} = \mathbf{r}\left(\mathbf{V}^{\mathrm{T}} - \mathbf{U} + \hat{\mathbf{s}}\right)^{-1} = \mathbf{r}\mathbf{A}^{-1} \tag{1}$$

In our case, the four elements in **b** represent the production processes of Figure 4.1 (animal husbandry, manure storage, crop production, feed storage). The nitrogen use efficiency (NUE_N) for the production process of the animal husbandry is quantified by:

$$NUE_N = 1/b_{\mathrm{husbandry}} \tag{2}$$

A detailed example of this procedure can be found in the supplementary material of Suh and Yee (2011).

4.2.5 Quantifying the effect of epistemic uncertainty on benchmarking

To quantify the effect of epistemic uncertainties of the input parameters on the benchmarking of farms based on their NUE_N, the distribution functions of the parameters need to be defined first. Subsequently, the input uncertainties are propagated through the NUE_N model.

Defining distribution functions

Each parameter in the NUE_N model was considered as an uncertain parameter, only the N-flow from crop production to feed storage and the N losses during manure storage were fixed. The N-flow from crop production was fixed, because it was based on the energy requirements of the herd. The N losses during manure storage were fixed, because they were based on

storage specific emission factors. All input parameters are assumed to be normally distributed. Fixation was assumed to be truncated normally distributed to avoid drawing negative numbers. The coefficient of variation (CV= σ/μ) described the epistemic uncertainty of the parameters and was based on Oenema et al. (2015) (Table 4.2). Based on the equation for the CV, the standard deviation was calculated per farm, because each farm had a different (i.e. farm specific) mean.

Quantifying the effect of epistemic uncertainty on benchmarking

The propagation of the uncertainties of the input parameters through the NUE_N model (equation 1) was done using Monte Carlo simulation and was performed for all farms in each cluster. The code for performing the uncertainty and global sensitivity analysis is available at: http://evelynegroen.github.io. From each distribution function (Table 4.1) a random value was drawn, and used to calculate the NUE_N. The output uncertainty was given by the variance:

$$\text{var}(NUE_N) = \frac{1}{n-1}\sum_{i=1}^{n}\left(NUE_{N_i} - \overline{NUE_N}\right)^2$$

Where the mean is given by: $\overline{NUE_N} = \frac{1}{n}\sum_i NUE_{N_i}$, for a sample size of *n*=5000. We performed a discernibility analysis (Heijungs & Kleijn, 2001) to determine if the input uncertainties had an effect on benchmarking. To determine if there was a significant difference between farms the farms within a cluster were pairwise compared for the results for each Monte Carlo run. This means that we counted how many times the NUE_N of one farm was better than another farm, expressed as a frequency. A significance level of 5% was chosen (Heijungs & Kleijn, 2001; Henriksson et al. 2015). This means, for example, that if farm A has a lower NUE_N than farm B in 630 out of 1000 runs, difference in NUE_N of the two farms was considered as not significant (63%>2.5%). But, if farm A had a lower NUE_N than farm C in 24 out of 1000 runs, than farm C was considered as significantly better than farm A (2.4%<2.5%).

4.2.6 Explaining output uncertainty for different farm typologies

To identify which input parameter contributed most to the output uncertainty within a specific farm cluster, a global sensitivity analysis was performed by calculating the squared standardized regression coefficients (S_j) as a measure for the sensitivity index (Saltelli et al., 2008; Groen, 2016):

$$S_j = \frac{\text{var}(p_j)}{\text{var}(NUE_N)}(b_j)^2$$

Where $\text{var}(p_j)$ gives the variance of each input parameter (p_j) based on Table 4.2 and b_j is equal to the regression coefficient.

Table 4.2 Description of the parameters and their epistemic uncertainty given by the relative uncertainty (CV), which was taken from Oenema et al. (2015).

Process	Type	Parameter		CV (%)	Remark
Crop production	Resource input	N-fixation (kg N)	Grassland area (ha)	5.0	
			Legume yield (kg/ha)	10	
			N-fixation (kg N/kg legume)	30	
		Deposition (kg N)	Farm area (ha)	5.0	
			N-deposition (kg N/ha)	17	
		Mineral fertilizer (kg N)	Mineral fertilizer (kg)	2.5	
			N-content mineral fertilizer (kg N/kg)	2.5	
		Stock change mineral fertilizer (kg N)	Stock change mineral fertilizer (kg)	7.5	
			N-content stock change mineral fertilizer (kg N/kg)	2.5	
	Export	Roughage (kg N)	Roughage (kg)	7.5	
			N-content roughage (kg N/kg)	7.5	
	Losses	N-losses crops (kg N)	n.a.	n.a	Function
Feed storage	Resource input	Roughage (kg N)	Roughage (kg)	7.5	
			N-content roughage (kg N/kg)	7.5	
		Stock change roughage (kg N)	n.a.	17	GEP[I]
Fertilizer storage	Resource input	Manure (kg N)	Manure (kg)	5.0	
			N-content manure (kg N/kg)	7.5	
		Stock change manure (kg N)	n.a.	22	GEP
	Losses	N emissions from manure storage (kg N)	n.a.		Fixed
		Animals (kg N)	Number of animals (-)	2.0	

Milk and animal production	Resource inputs		Life-weight per animal (kg)		n.a.
			N-content per animal (kg N/kg)	5.0	
		Stock change animals (kg N)	n.a.	5.6^8	GEP
		Concentrates (kg N)	Concentrates (kg)	2.5	
			N-content concentrates (kg N/kg)	2.5	
		Stock change concentrates (kg N)	n.a.	11	GEP
	Final use	Milk (kg N)	Milk (kg)	1.0	
			N-content milk (kg N/kg)	2.0	
		Animals (kg N)	Number of animals (-)	2.0	
			Life-weight per animal (kg)		n.a.
			N-content animal (kg N/kg)	5.0	
	Export	Manure (kg N)	Manure (kg)	5.0	
			N-content manure (kg N/kg)	7.5	

[I]GEP: Gaussian error propagation is used to determine the CV of parameters when there is a lack of information to separate the N-content from the items in the stock change and therefore only the kg N of stock change is available (e.g., roughage can include different items with different N contents). Details on the method can be found in (Heijungs and Lenzen, 2014).

4.3 Results

4.3.1 Farm clusters

Two homogeneous groups of farms, i.e. farm clusters, were derived from the cluster analysis. Farms in the first group, further referred to as farms in cluster 1, are located on sandy soils and relatively intensive in terms of milk production per cow and per hectare (Table 4.3). The number of grazing hours is low, whereas the amount of purchased concentrates per cow per year is high relative to the farms in the other cluster. Farms in cluster 2 are located on loam soils, and are less intensive when compared to farms in cluster 1. The number of grazing hours is higher, whereas the amount of concentrates per cow per year is lower than on farms in cluster 1. The average NUE_N of farms in cluster 1 is 43%, and for farms in cluster 2 this is 26%. The difference in NUE_N between the two clusters result from a combination of all 5 characteristics that specify the group of farms in each cluster (Table 4.3).

Table 4.3 Results of the cluster analysis, showing the farm characteristics for 15 farms in cluster 1 and 17 farms in cluster 2, given by the mean (standard deviation) of each characteristic or a categorical characteristic per cluster.

Characteristics[1]	Unit	Cluster 1	Cluster 2
Soil type	n.a.	Sandy	Loam
Milk production	kg milk cow^{-1} $year^{-1}$	8519 (854)	6956 (878)
Milk production	kg milk cow^{-1} ha^{-1}	15970 (5108)	9623 (3792)
Grazing hours	hours cow^{-1} $year^{-1}$	1115 (1099)	4393 (1175)
Concentrate use	kg $cow^{-1}year^{-1}$	1719 (499)	770 (207)
NUE_N	%	43 (10)	26 (12)

[1] Characteristics of these two clusters are significantly different (p<0.05). The order of importance of the characteristics in determining the final clusters are: grazing hours > concentrate use > milk production per cow > soil type > milk production per ha.

4.3.2 The effect of epistemic uncertainties on benchmarking

For each farm, the input uncertainties of Table 4.2 were propagated through the NUE_N model (equation 1 and 2). For each farm in both clusters, a median and a variance were derived (Figure 4.2, cluster 1; Figure 4.3, cluster 2). Results show that each cluster has one outlier: farm 1 in cluster 1 and farm 2 in cluster 2. For farm 1 in cluster 1, the output of manure exceeds the input of fertilizer. Because we subtracted manure output from fertilizer inputs, the input of fertilizer was set to 0. This leads to the high NUE_N of this farm. Farm 2 in cluster 2 is an organic farm with only grassland and no cropland. The imported feed inputs are low, and there is no input of synthetic fertilizer. Due to the low N inputs and high N outputs of the farm, it has a high NUE_N.

Figure 4.2 Box plot of NUE_N for the 15 farms in cluster 1. The horizontal line in each box gives the median, the box gives the 25-75% interval, and the plusses are realizations that appear outside the 10-90% interval.

The results of the discernibility analysis for cluster 1 can be found in Table 4.4. For example, farm 5 had a lower NUE_N than farm 1, and a higher NUE_N than the other farms, except when compared to farm 8 and farm 14. In case of farm 6, only 52% of the Monte Carlo runs show a higher NUE_N than farm 3, meaning their performance is almost indistinguishable taking the epistemic uncertainties of the input parameters into account.

For farm 1, approximately 4% of the Monte Carlo runs resulted in a negative value for N losses of crop production. This is explained by the importance of deposition as an N input on this farm, and the large uncertainty of this parameter (CV=17%; Table 4.2). The negative values, therefore, are more likely related to the uncertainty of deposition, than to display a realistic model outcome. The drawings from the Monte Carlo simulation that included a negative value for N losses of crop production, therefore, were removed from the analysis.

Applying the 5% significance level, results show that farm 1 is most efficient when taking the epistemic uncertainty of the input parameters into account, followed by farm 5, which is only

not significantly better than farm 8 and 14. The two least efficient farms are farm 3 and 6. The NUE$_N$ of the other farms turned out to be very similar (Table 4.4).

Table 4.4 Results of discernibility analysis for cluster 1 based on pairwise comparing Monte Carlo runs between farms. The column and row numbers 1 to 15 represent the 15 farms. The percentages show how often a farm (row) has a higher NUE$_N$ than another farm (column). When α-value of 0.05 is applied, values between 2.5% and 97.5% indicate that the NUE$_N$ of the farms are no longer considered as significantly different. The significant different farms are given by the bold-printed percentages.

%	1*	2	3	4	5	6	7	8	9	10	11	12	13	14	15
1*		**100**	**100**	**100**	**100**	**100**	**100**	**100**	**100**	**100**	**100**	**100**	**100**	**100**	**100**
2	**0**		**100**	5	**0**	**100**	25	5	78	47	52	19	54	7	89
3	**0**	**0**		**0**	**0**	48	**0**	**0**	**0**	**0**	**0**	**0**	**0**	**0**	**0**
4	**0**	95	**100**		**2**	**100**	85	43	**99**	92	88	70	97	42	**100**
5	**0**	**100**	**100**	**98**		**100**	**100**	95	**100**	**100**	99	**98**	**100**	91	**100**
6	**0**	**0**	52	**0**	**0**		**0**	**0**	**0**	**0**	**0**	**0**	**0**	**0**	**0**
7	**0**	75	**100**	15	**0**	**100**		13	93	69	71	37	79	16	97
8	**0**	95	**100**	57	5	**100**	87		**99**	93	90	73	97	49	**100**
9	**0**	22	**100**	1	**0**	**100**	7	**1**		22	31	6	24	**1**	67
10	**0**	53	**100**	8	**0**	**100**	31	7	78		55	23	57	9	88
11	**0**	48	**100**	12	1	**100**	29	10	69	45		23	50	11	81
12	**0**	81	**100**	30	**2**	**100**	63	27	94	77	77		84	28	**97**
13	**0**	46	**100**	3	**0**	**100**	21	3	76	43	50	16		5	88
14	**0**	93	**100**	58	9	**100**	84	51	**99**	91	89	72	95		**99**
15	**0**	11	**100**	**0**	**0**	**100**	3	**0**	33	12	19	3	12	**1**	

* Approximately 4% of the Monte Carlo runs were excluded from the analysis due to unrealistic model outcomes.
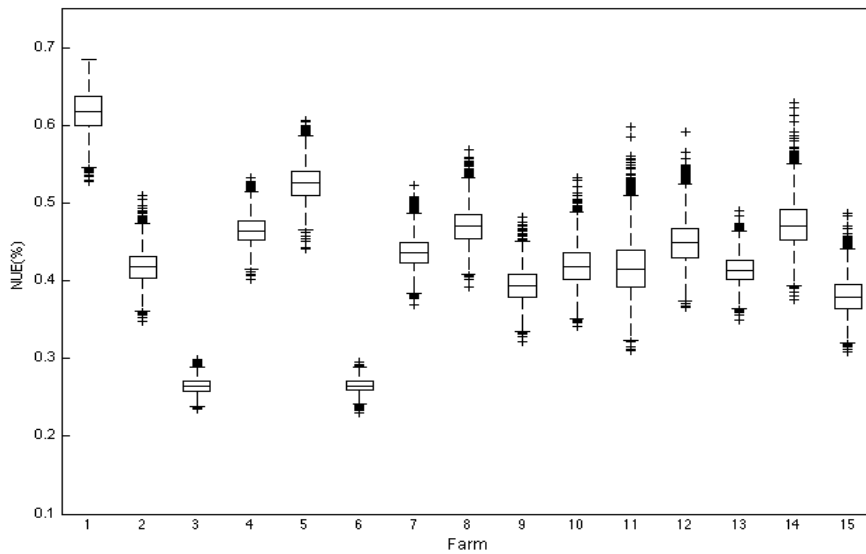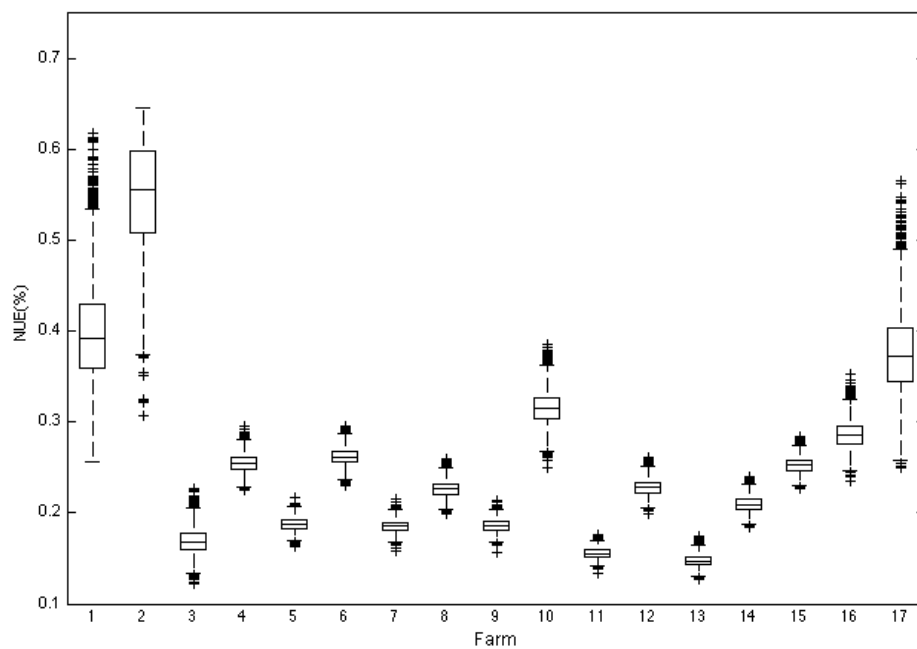
Figure 4.3 Box plot of NUE_N for the 17 farms in cluster 2. The horizontal line in each box gives the median, the box gives the 25-75% interval, and the plusses are realizations that appear outside the 10-90% interval.

The results of the discernibility analysis for cluster 2 are found in Table 4.5. For farm 2, approximately 46% of the Monte Carlo runs resulted in a negative value for N losses of crop production. This is explained by the importance of N fixation on this farm, in combination with a relatively large uncertainty of this parameter (CV=30%). Because we analysed quite intensive and productive farms, these outcomes are more likely to result from the high CV than to display a realistic situation. Similar to farm 1 in cluster 1, negative values were assumed to display an unrealistic model outcome. Because of the high percentage of unrealistic model outcomes, it was decided to remove farm 2 from further analysis. The large number of unrealistic model outcomes illustrates the need to reduce CVs by improving measurements on farms.

Applying the 5% significance level, results show that farm 1 is most efficient (only not significantly higher than farm 10 and 17). Of the 120 farm comparisons, 17% is significantly

different (Table 4.5). Contrary to the first cluster, including the epistemic uncertainties still allowed for some kind of ranking, although most farms overlapped with at least two other farms.

Table 4.5 Results of discernibility analysis for cluster 2 based on pairwise comparing Monte Carlo runs between farms. The column and row numbers 1 to 17 represent the 16 farms*. The percentages show how often a farm (row) has a higher NUE_N than another farm (column). When α-value of 0.05 is applied, values between 2.5% and 97.5% indicate that the NUE_N of the farms are no longer considered as significantly different. The significant different farms are given by the bold-printed percentages.

%	1	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17
1		**100**	**100**	**100**	**100**	**100**	**100**	**100**	95	**100**	**100**	**100**	**100**	**100**	**99**	62
3	**0**		**0**	12	**0**	14	**0**	14	**0**	82	**0**	92	**1**	**0**	**0**	**0**
4	**0**	**100**		**100**	31	**100**	**99**	**100**	**0**	**100**	**98**	**100**	**100**	57	4	**0**
5	**0**	88	**0**		**0**	59	**0**	59	**0**	**100**	**0**	**100**	**2**	**0**	**0**	**0**
6	**0**	**100**	69	**100**		**100**	**100**	**100**	**0**	**100**	**100**	**100**	**100**	76	9	**0**
7	**0**	86	**0**	41	**0**		**0**	49	**0**	**100**	**0**	**100**	**1**	**0**	**0**	**0**
8	**0**	**100**	**1**	**100**	**0**	**100**		**100**	**0**	**100**	45	**100**	93	**2**	**0**	**0**
9	**0**	86	**0**	41	**0**	51	**0**		**0**	**100**	**0**	**100**	**1**	**0**	**0**	**0**
10	5	**100**	**100**	**100**	**100**	**100**	**100**	**100**		**100**	**100**	**100**	**100**	**100**	90	8
11	**0**	18	**0**	**0**	**0**	**0**	**0**	**0**	**0**		**0**	83	**0**	**0**	**0**	**0**
12	**0**	**100**	**2**	**100**	**0**	**100**	55	**100**	**0**	**100**		**100**	94	**2**	**0**	**0**
13	**0**	8	**0**	**0**	**0**	**0**	**0**	**0**	**0**	17	**0**		**0**	**0**	**0**	**0**
14	**0**	**99**	**0**	**98**	**0**	**99**	7	**99**	**0**	**100**	6	**100**		**0**	**0**	**0**
15	**0**	**100**	43	**100**	24	**100**	**99**	**100**	**0**	**100**	**98**	**100**	**100**		**2**	**0**
16	**1**	**100**	96	**100**	91	**100**	**100**	**100**	11	**100**	**100**	**100**	**100**	**98**		**1**
17	37	**100**	**100**	**100**	**100**	**100**	**100**	**100**	92	**100**	**100**	**100**	**100**	**100**	**99**	

* Approximately 46% of the Monte Carlo runs were excluded from the analysis due to unrealistic model outcomes of farm 2, therefore, this farm was excluded from further analysis.

4.3.3. Explaining the output variance

The global sensitivity analyses shows how much of the output variance can be explained by the variance of the individual input parameters. The results of the global sensitivity analysis can be found in Figure 4.4 (cluster 1) and Figure 4.5 (cluster 2).

Parameter (%)	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
Input animals		0		0				0				0			0
Input concentrates	30	1	29	8	4	39	30	15	16	10	4	3	9	16	8
Input roughage	0		17	0	3		0	49	25	51	20	2	28	22	30
Input min. fertilizer	10	8	14	11	7	15	16		3	1	0	20	7	2	7
Input manure		2							1	0				1	
Deposition	21	4	4	7	3	6	7	15	7	9	13	16	30	9	7
Fixation									0		52			1	
SC animal	0	0	0	0	0	0	0		0	0	0	0	0	0	0
SC concentrates				0	0		0		0	0	0	0	0	0	0
SC roughage									0	1	2	1	1	4	2
SC min. fertilizer									0	1		31		0	
SC. manure									2	2	1		3	3	2
Output animal	1	0	5	1	0	4	1	1	2	1	1	1	2	1	1
Output milk	13	3	29	8	5	35	20	16	13	11	6	5	16	10	11
Output roughage	4	81	0	63	77							19	3		
Output manure	35	0	4				26		29	12	2		2	30	33
Explained variance	n/a	100	100	99	99	100	100	97	99	100	100	100	100	100	100

Figure 4.4 Sensitivity indices (S_j) for each input parameter, explaining how much each parameter contributes to the output variance for each farm in cluster 1: the darker a cell, the higher the contribution. SC: stock change. An empty cell means that these parameters were zero for that farm; 0% means that this parameter contributed 0% to the output variance; n/a not applicable, approximately 4% of the Monte Carlo runs were excluded from the analysis due to unrealistic model outcomes, therefore the partial variances were not considered independent and could not be added.

Results show that in case of cluster 1, the input of concentrates, roughage, mineral fertilizer, and deposition, and the output of milk, roughage, and manure explain most of the output variance. Input of animals and manure, stock change of each of the inputs, and output of animals did not show up as important explanatory parameters in any of the farms, except for stock change of mineral fertilizer for farm 12. Further analysis showed that both the quantity as well as the N content of each parameter is approximately equally important in terms of their contribution to the output variance.

Parameter (%)	1	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17
Input animals		0										12				0
Input concentrates	0	0	0	2	4	2	2	2	0	3	2	0	1	13	0	0
Input roughage		3	0	8	0	0	0	0	1	1		1	0	2		6
Input min. fertilizer		1	70	45	54	44	66	61	12	55	57	45	60	38	59	
Input manure	3	8														
Deposition	9	3	2	1	5	2	1	1	1	1	3	1	3	5	1	14
Fixation	82	78														68
SC animal	0	0	0	2	0	1	0	0	0	0	1	1	0	0	0	0
SC concentrates																0
SC roughage																1
SC min. fertilizer						13			10	0			0		15	
SC. manure																
Output animal	1	2	10	21	12	13	6	11	1	15	14	25	23	13	2	1
Output milk	2	5	17	21	25	27	25	24	2	22	21	15	14	28	8	1
Output roughage									73						11	8
Explained variance	97	99	99	100	100	100	100	98	100	99	97	100	100	99	97	99

Figure 4.5 Sensitivity indices (S_j) for each input parameter, explaining how much each parameter contributes to the output variance for each farm in cluster 2: the darker a cell, the higher the contribution. SC: stock change. An empty cell means that these parameters were to zero for that farm; 0% means that this parameter contributed 0% to the output variance. Output of manure is not included because none of the farms in cluster 2 exported manure. Farm 2 was excluded from the global sensitivity analysis.

Figure 4.5 shows that in case of cluster 2, for most of the farms the input of mineral fertilizer and the output of milk and animals explain most of the output variance. For a few farms, the most important parameter in terms of contribution to the output variance is fixation (input); while for one farm, this is the output of roughage. Input of concentrates, roughage and manure and stock change of animals did not show up as important explanatory parameters in any of the farms in cluster 2.

4.3.4 Effect of decreasing uncertainty on benchmarking

To analyse if decreasing epistemic uncertainty can improve benchmarking, we reduced the uncertainty of the most important input parameters and reran the discernibility analysis. For cluster 1, the input uncertainty was reduced to 1% for: input of concentrates, roughage, mineral fertilizer, deposition and the output of milk, roughage, and manure (Figure 4.4). For cluster 2,

the input uncertainty was reduced to 1% for: input of mineral fertilizer, and the output of milk and animals (Figure 4.5). Table 4.6 shows how many pairwise comparisons were made in both cluster, and how many were significantly different, before and after reducing input uncertainty. Results show that reducing the uncertainty of the most important input parameters based on the global sensitivity analysis, improved the ability to find significant differences between the NUE_N of the farms in both clusters. Benchmarking, therefore, can be improved when input uncertainties are reduced, especially for the farms in the first cluster.

Table 4.6 Effect of decreasing the input uncertainties of the most important parameters to 1%, for both clusters.

	Cluster 1	Cluster 2
Total number of pairwise comparisons	105	120
Significantly different farms before reducing input uncertainty	39 (37%)	99 (83%)
Significantly different farms after reducing input uncertainty	76 (72%)	104 (87%)

4.4 Discussion

This study builds on, and extends the principles regarding epistemic uncertainty of nitrogen flows on dairy farms presented by Oenema et al. (2015). Although we used the same coefficients of variations of input parameters, results of our study and Oenema et al. (2015) show important differences. Based on our analysis, input of concentrates and roughage, and output of milk and roughage explain most of the output variance in cluster 1. Input of mineral fertilizer and fixation, and output of animals and milk explain most of the output variance in cluster 2. Oenema et al. (2015), however, concluded that N fixation, atmospheric deposition and stock changes of roughage and manure explain most of the output variance when determining the N surplus of dairy farms. Differences between our study and Oenema et al. (2015) can be explained by two reasons. First, the characteristics of the farms were different. In general, Oenema et al. (2015) included farms with a lower input of feed, but a higher stock change of roughage, and a higher N input through fixation compared to the farms in our study. Uncertainties related to stock changes of feed are higher than uncertainties related to input of

feed, whereas uncertainties related to N fixation is highest among all N flows. Second, Oenema et al. (2015) used a different approach to determine N intake during grazing. In our study, N intake from grazing and on-farm roughage production was fixed based on feed requirements and the baseline values of input of purchased feed (see supplementary materials). Oenema et al. (2015) changed the N-intake from grazing with a change in roughage and concentrate intake, which consequently influenced the importance of feed parameters. The contribution of the input of feed to the output variance was therefore found to be lower in Oenema et al. (2015) than in our study.

Dairy farms in Europe show different decisive characteristics. For example, most farms in the Netherlands have a high stocking density because land resources are limited. The main N inputs on these farms are through purchased concentrates and roughages. In Ireland, however, most farms are grass-based extensive farms. The main N inputs on these farms are through purchased mineral fertilizer and N fixation. Comparing NUE of Dutch and Irish farms can lead to biased conclusions because of inherent differences between systems. Clustering of farms into groups with similar decisive characteristics, therefore, is a prerequisite for benchmarking the NUE of farms and facilitates the identification of major parameters. When comparing results of the global sensitivity analysis between the two farm clusters, for example, input of feed and output of roughage show up to be most important in case of cluster 1, whereas the input of mineral fertilizer (or fixation) is most important in case of cluster 2. Results show that the importance of parameters can vary between farm types (clusters). Methods to improve benchmarking of farms, therefore, should account for differences in decisive characteristics. The method presented in this study, can contribute to more solid conclusions regarding the performance of farms in terms of their NUE.

In this study, we used a matrix-based approach to assess NUE. The advantage of this approach is that it facilitates the uncertainty and sensitivity analysis. All input parameters are sampled at the same time, and are subsequently used to calculate the internal flows for each Monte Carlo run. Another advantage of the matrix-based approach is that it is easy to extend

the system boundary beyond the farm: production of crops can be easily incorporated as additional production flows.

Several methodological limitations could have affected the results of this study. The first limitation is the choice of the parameter distribution function and, in link to our choice of Gaussian distribution, of CV values. A Gaussian or normal distribution represents a symmetrical uncertainty range which seems correct in case of (most) measurement errors. Future studies, however, could use parameter specific distribution functions to improve the impact assessment. The CVs we used were based on Oenema et al. (2015), focussing on Dutch pilot commercial farms only. Farms in our study are from different countries in Europe. Results of the uncertainty and global sensitivity analysis might have been different if country specific coefficients of variation were applied, but such information was not available. In addition, we used farm data from the year 2010, which might not hold for any year. However, since a similar measurement error over years can be expected, we do not expect a big change in the CVs, but mainly in the mean values of the N-flows on the farms. In general, it takes quite a big change in the CV to influence the result of the sensitivity analysis as seen in Figure 4.4 and 4.5. In addition, only drastically decreasing the CVs of the most important parameters (e.g. from 30% to 1%) influenced the number of significantly different farms (mainly in the first cluster; Table 4.6). Differences in the CVs because of yearly variations, therefore, are not expected to influence the results. Nevertheless, the methodological procedure that was presented in this study can be used to assess the impact of epistemic uncertainty on different farms and based on different CVs. Results show that to benchmark the NUE of farms, epistemic uncertainty of input parameters has to be reduced.

Secondly, changes in soil N-stock were not considered in this study due to data limitations. Assessing changes in soil N stock at the farm level is difficult but can significantly improve interpretation of nutrient balance results (Godinot et al., 2014).

Thirdly, uncertainty related to on-farm crop and grass production was not included in the model, because this was estimated based on the energy requirements of the dairy herd and the energy in purchased feed and stock changes of feed. Incorporating uncertainty of crop production would increase the uncertainty of the model output. This would mainly affect the farms relying more heavily on on-farm produced roughage, such as the farms in cluster 2. Considering that input of mineral fertilizer or fixation and output of N via milk an animals are the main explanatory factors of the uncertainty around the NUE of these farms, we do not expect that the additional uncertainty of on-farm crop (grass) production would influence the results much.

Fourthly, to prevent purchase-resale bias (Godinot et al., 2014), the output of roughage was subtracted from the input of roughage, and the output of manure was subtracted from the input of fertilizers. As a result, exported manure is valued for its fertilizer capacity similar to (synthetic) fertilizer inputs. The disadvantage of the approach is that the output of manure results in an artificial reduction of fertilizer input, while an actual reduction should form the basis for ecological intensification and an improved NUE. The importance of manure output and the impact of these methodological choices should be addressed when benchmarking the NUE of dairy farms.

Fifthly, clustering of farms was based on 5 characteristics reflecting physical and long term strategic decisions. In practice, farming systems are much more complex than we considered in our study. Including other (unmanageable) factors that affect NUE could influence the clustering of farms and hence, the benchmarking of those farms. Nevertheless, this study is a first step towards improving benchmarking farms based on their NUE. Results emphasize the need to benchmark NUE by comparing farms with similar decisive characteristics, and that the importance of parameters that contribute to the uncertainty of the NUE results differ among farm types.

Sixthly, this study focused on NUE_N at farm level. Nitrogen losses related to the production of purchased feed and fertilizers were not considered. It should be kept in mind that, as a result, the NUE of a farm increases with a decrease in self-sufficiency. This approach, therefore, can contribute to biased conclusions and problem swapping, when on-farm nutrient losses related to feed production are reduced at the expense of off-farm losses. Furthermore, NUE provides insight into the efficiency of production rather than into the environmental impact related to nutrient losses. To gain insight into the impact of losses, information on nutrient losses per hectare should be combined with site specific knowledge of local eco-systems. In addition to this, it should be noted that the results of this study are limited to benchmarking the NUE_N of specialized dairy farms in Europe. For another indicator or another set of farms, the impact of uncertainty on benchmarking the environmental performance of dairy production could be different.

Reducing epistemic uncertainty and benchmarking NUE of farms with similar decisive characteristics can contribute to the identification of improvement options. Based on the variability between farms within a cluster, farm specific management options can be identified. Evaluating the (causes of) variability between farms within a farm cluster, therefore, can be a next step for further improving the NUE of farms.

4.5 Conclusion

Benchmarking the NUE of dairy farms requires an approach that accounts for differences in major decisive characteristics among farms, and for the impact of epistemic uncertainties of input parameters. The parameters that are most important in terms of epistemic uncertainty (i.e., explain most of the output variance), however, can vary among farm types. Clustering farms based on their main characteristics and understanding and reducing the impact of epistemic uncertainty of major parameters can significantly improve benchmarking results. The method presented in this study, therefore, can contribute to more solid conclusions regarding the performance of farms in terms of their nutrient use efficiency.

Acknowledgements

Supplementary material

S1. Example of nutrient balances using a matrix-based approach

Figure S.1 shows a nutrient balance of a dairy farm. The total input of N (in kg) is by the dark grey boxes, the light grey boxes give the output. The uptake from the stock change of concentrates (C) is subtracted from the total input of concentrates, because it is considered as an input to the farm. The addition to the stock of animals (A) is considered as a useful output of the farm, and is therefore added to the output of animals. The amount of produced manure that is exported is subtracted from the input of fertilizers (i.e., mineral fertilizer and manure). Hence, manure export is considered to offset a farm's fertilizer input. However, if export of manure exceeded the input of fertilizer, it was considered as a loss. The nutrient use efficiency of nitrogen (NUE_N) on this farm is equal to 44%.
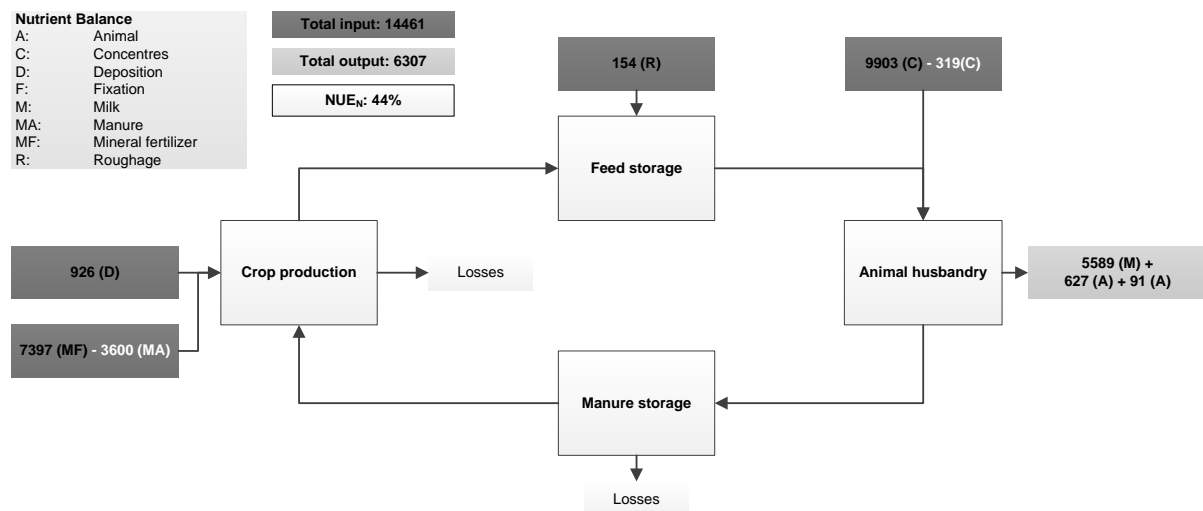

Figure S.1. Nutrient balance of a dairy farm (kg N). The total input of N (in kg) is by the dark grey boxes, the light grey boxes give the output.

Figure S.2 shows the same farm, with similar assumptions as described above, but in this case the NUE_N is quantified using a matrix-based approach as described by Suh and Yee (2011). When the same assumptions are applied, the matrix-based approach results in the same efficiency of 44%. However, a difference is that the efficiency of the production processes (i.e. crop production, feed storage, animal husbandry and fertilizer storage), are given with respect

to the whole farm, not of the individual processes, in case of the matrix-based approach. For example, the efficiencies of feed storage and animal husbandry are 100%, when looking at the individual processes, because there are no losses. The matrix-based NUE_N however, incorporates the losses that occurred elsewhere on the farm in the efficiency of the individual production processes. For example, the efficiency of feed storage equals 28%, which is determined by the efficiency of the previous processes, such as crop production, and input of roughage. The efficiency of animal husbandry equals the efficiency of the farming system (44%), because it comprises the efficiency of all former processes.
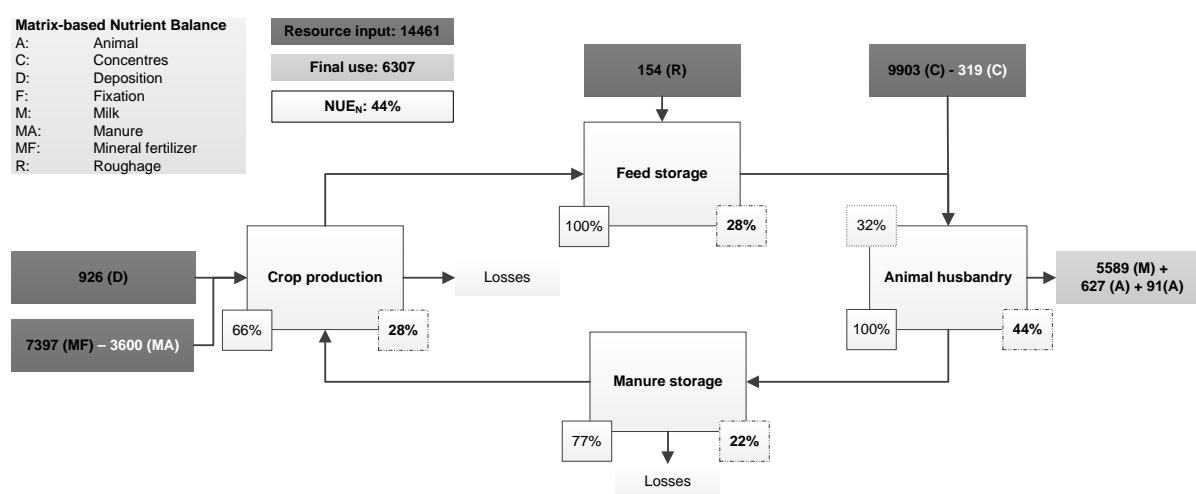

Figure S.2. Nutrient balance of a dairy farm, in kg N. The total input of N (in kg) is by the dark grey boxes, the light grey boxes give the output. The efficiencies in the bottom right corner display the efficiencies quantified with the matrix based nutrient balance, the efficiencies in the bottom left corner display the efficiencies quantified looking only at the in- and outputs of that specific production system. If an efficiency of 100% is give, it means that there are no losses at that process. The efficiency of 32% in the top left corner (animal husbandry) displays the efficiency when the export of manure is considered as a loss.

Figure S.3 shows for the same farm, what happens if the export of manure is no longer subtracted from the input of fertilizer, but considered as an individual export of the farm. In that case, for the matrix based nutrient balance, there are now two efficiencies, one describing the useful output we considered earlier (milk and animal) and one describing the production of manure.
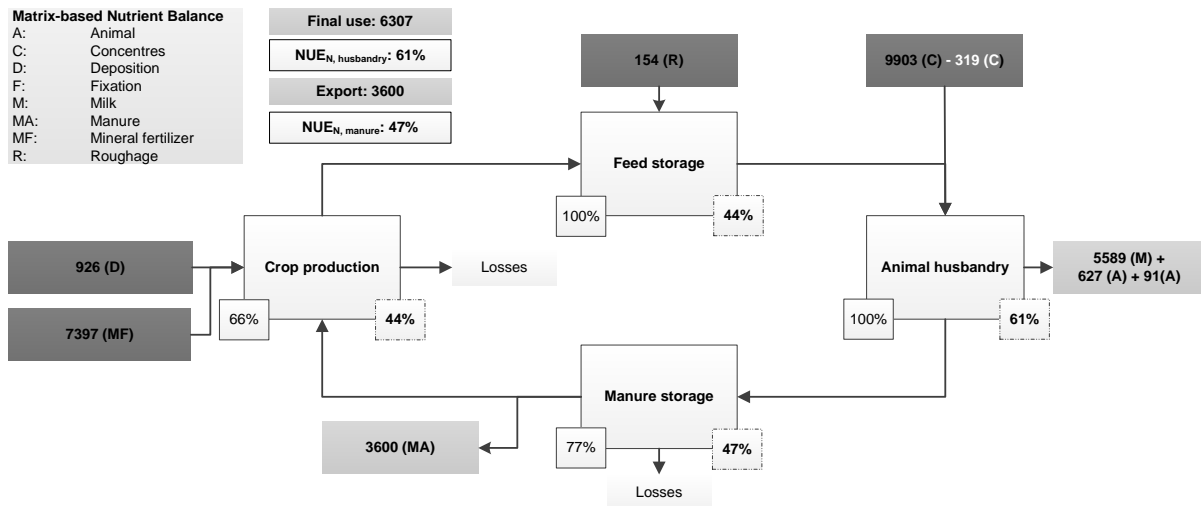
Figure S.3. Nutrient balance of a dairy farm, in kg N, when export of manure is considered as an output. The total input of N (in kg) is by the dark grey boxes, the light grey boxes give the output. The efficiencies in the bottom right corner display the efficiencies quantified with the matrix based nutrient balance, the efficiencies in the bottom left corner display the efficiencies quantified looking only at the in- and outputs of that specific production system. If an efficiency of 100% is given, it means that there are no losses at that process.

In this paper, we always subtracted manure from the input of fertilizer, including both mineral fertilizer and manure. However, some farms exported roughage. Based on the same principles as described above (i.e., considering the export of manure as an output), export of roughage was considered as an output. This resulted in a slightly different NUE_N than what would have been calculated using the normal nutrient balance approach, but differences were minor (1-2%).

S2. Calculating on-farm grass and crop production

On-farm grass and crop production is estimated based on the energy requirement of the herd, minus energy contained in purchased feed and stock change of the feed.

Step 1: The amount of consumed concentrates, consumed purchased roughage and consumed purchased by-products are determined based on farm data regarding the input and stock changes of these feed products. Data were obtained from Dairyman (2010).

Step 2: The energy content in consumed concentrates, consumed purchased roughage and consumed purchased by-products based on CVB (2008) are multiplied with the amount of each feed product to determine energy intake from purchased feed and stock changes of the feed.

Step 3: The total energy requirement of the dairy herd is calculated based on the energy required for milk production (dairy cows), maintenance, grazing, growth (young stock and heifers) and pregnancy according to CVB (2008).

Step 4: Energy from on-farm grass (crop) production is derived from subtracting energy of consumed concentrates, energy of consumed purchased roughage, energy of consumed purchased by-products from the total energy requirement of the dairy herd.

Step 5: Based on the hectares of crop and grassland on the farm, an average grass (crop) yield per ha, and the energy content of the grass (crop) products, the relative share of each of the grass (crop) products produced on the farm is determined. In this step, we also included harvest losses.

Step 6: On-farm grass and crop production is determined by multiplying the relative share of each grass (crop) product produced on the farm (Step 5) with the energy requirements to be fulfilled by on-farm grass (crop) production (Step 4).

References

CVB, 2008. Chemische samenstelling en nutritionele waarden van voedermiddelen Productschap Diervoeder, Den Haag, The Netherlands, ISSN 1567–8679

Dairyman,2010. < http://www.interregdairyman.eu/en/dairyman.htm> Last access on 14-01-2016.

S3. Calculation rules nutrient balances

Table S.1. Meaning of the input parameters (*x*) in Table S.2 and S.3. The bold-printed item *x_loss_3* is a function of the other input parameter in the NUE model.

	Input	Stock change	Losses	Output
1a. Animal	*x_in_1*	*x_stock_1*	*n.a.*	*x_out_1a*
1b. Milk				*x_out_1b*
2. Concentrates	*x_in_2*	*x_stock_2*	*n.a.*	
3. Roughage	*x_in_3*	*x_stock_3*	***x_loss_3***	*x_out_3*
4. Mineral fertilizer	*x_in_4*	*x_stock_4*		
5. Organic fertilizer	*x_in_5*	*x_stock_5*	*x_loss_5*	*x_out_5*
6. Deposition	*x_in_6*		*n.a.*	
7. Fixation	*x_in_7*		*n.a.*	

Table S.2. Calculation rules for the intermediate flow. The bold-printed items are functions of other parameters.

Intermediate flows:		Function:
Flow from crop production to feed storage	*x_int_A =*	Fixed
Flow from feed storage to animal husbandry	***x_int_B =***	*x_int_A+ x_in_3 - x_stock_3*
Flow from animal husbandry to manure storage	***x_int_C =***	***x_int_B** + x_in_1 - x_stock_1 - x_out_1b - x_out_1a + x_in_2 - x_stock_2*
Flow from manure storage to crop production	***x_int_D =***	***x_int_C** + x_in_ 5 - x_loss_5 - x_out_5 - x_stock_5*
Losses manure storage	***x_loss_3 =***	***x_int_D** + x_in_6 + x_in_7 - **x_int_A** - x_out_3 + x_in_4 - x_stock_4*

Table S.3. These equations can be used to make sure that the mass balances hold for each process.

Mass balances	In:	Out:	Losses
Fertilizer storage balance	*x_int_C + x_in_5*	*x_int_D + x_out_5 + x_stock_5*	*x_loss_5*
Crop or soil balance	*x_int_D + x_in_6 + x_in_7 + x_in_4 - x_stock_4*	*x_int_A + x_out_3*	***x_loss_3***
Feed storage balance	*x_int_A + x_in_3*	*x_int_B + x_stock_3*	*n.a.*
Animal husbandry or herd balance	*x_int_B + x_in_1 + x_in_2 - x_stock_2*	*x_int_C + x_out_1b + x_out_1a + x_stock_1*	*n.a.*

4.6 References

Aarts, H.F.M., Biewinga, E.E., Van Keulen, H., 1992. Dairy farming systems based on efficient nutrient management. Neth. J. Agric. Sci. 40, 285–299.

Camp, R.C.. 1989. Benchmarking: The Search for Industry Best Practices that Lead to Superior Performance, ASQC Quality Press, Milwaukee, WI.

Chiu T., Fang D., Chen J., Wang Y., Jeris C., 2001. A robust and scalable clustering algorithm for mixed type attributes in large database environment. In: Proceedings of the 7th ACM SIGKDD international conference in knowledge discovery and data mining, Association for Computing Machinery, San Francisco, CA, 263–268.

Daatselaar, C.H., Reijs, J.R., Oenema, J., Doomewaard, G.J., Aarts, H.F., 2015. Variation in nutrient use efficiencies on Dutch dairy farms. J. Sci. Food Agric. 95,15.

Dairyman, 2010. http://www.interregdairyman.eu/en/dairyman.htm Last access on 14-01-2016.

European Environment Agency (EEA), 2013. EMEP/EEA Air Pollutant Emission Inventory Guidebook. http://www.eea.europa.eu/publications/emep-eea-guidebook-2013

European Union (EU), 2006. (Good Agricultural Practice for Protection of Waters) Regulations. S.I. No. 378 of 2006. Government Publications Office, Stationery Office.

Formann, A.K., 1984. Die Latent-Class-Analyse: Einführung in die Theorie und Anwendung. Beltz, Weinheim.

Godinot, O., Carof, M., Vertes, F., Leterme, P., 2014. SyNE: An improved indicator to assess nitrogen efficiency of farming systems. Agric. Syst. 127, 41-52.

Gourley, C.J.P., Aarons, S.R., Powell, J.M., 2012. Nitrogen use efficiency and manure management practices in contrasting dairy production systems. Agric. Ecosyst. Environ, 147, 73–81.

Groen, E.A., Bokkers, E.A.M. Heijung, R. and de Boer, I.J.M. (2016). Methods for Global Sensitivity Analysis in Life Cycle Assessment. Int. J. Life. Cycle. Assess. doi:10.1007/s11367-016-1217-3

Heijungs, R., Kleijn, R., 2001. Numerical approach towards life cycle interpretation. Int. J. Life. Cycle. Assess. 6, 141 – 148.

Heijungs, R., Lenzen, M., 2014. Error propagation methods for LCA—a comparison. Int J Life Cycle Assess. 19,1445–1461.

Henriksson, P.J.G., Heijungs, R., Dao, H.M., Phan, L.T., de Snoo, G.R., Guinee, J.B., 2015 Product Carbon Footprints and Their Uncertainties in Comparative Decision Contexts. PLoS One.1-11.

Mihailescu, E., Ryan, W., Murphy, P.N.C., Casey, I.A., Humphreys, J., 2015. Economic impacts of nitrogen and phosphorus efficiency on nineteen intensive grass-based dairy farms in the South of Ireland. Agric. Syst. 132, 121–132.

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.

Mulier, A. , Hofman, G. , Baecke, E. , Carlier, L. , De Brabander, D. , De Groote, G. , 2003. A methodology for the calculation of farm level nitrogen and phosphorus balances in Flemish agriculture. Eur. J. Agron, 20, 45–51.

Nevens, F., Verbruggen, I. , Reheul, D., Hofman, G., 2006. Farm gate nitrogen surpluses and nitrogen use efficiency of specialized dairy farms in Flanders: Evolution and future goals. Agric. Syst. 88, 142–155.

Oenema, J., Burgers, S., van Keulen, H., van Ittersum, M., 2015. Stochastic uncertainty and sensitivities of nitrogen flows on dairy farms in The Netherlands. Agric. Syst. 137, 126-138.

Oenema, O., Kros, H., deVries, W., 2003. Approaches and uncertainties in nutrient budgets: implications for nutrient management and environmental policies. Eur. J. Agric. 20, 3–16.

Phuong, H.N., Friggens, N.C., de Boer, I.J.M., Schmidely, P., 2013. Factors affecting energy and nitrogen efficiency of dairy cows: A meta-analysis. J. Dairy Sci. 96, 7245–7259.

Powell, J.M., Gourley, C.J.P., Rotz, C.A., Weaver, D.M., 2010. Nitrogen use efficiency: A potential performance indicator and policy tool for dairy farms. Environ. Sci. Policy. 13, 217-228

Roberts, T.L., 2008. Improving nutrient use efficiency. Turk. J. Agric. For. 32, 177–182.

Ryan, W., Hennessy, D., Murphy, J.J., Boland, T.M., Shalloo. L., 2011. A model of nitrogen efficiency in contrasting grass-based dairy systems. J. Dairy Sci. 94, 1032–1044.

Saltelli, A., Ratto, M., Andres, T., Corporation, E., 2008. Global sensitivity analysis: the primer. Wiley Online Library.

Schröder, J.J., Aarts, H.F.M., Ten Berge, H.F.M., Van Keulen, H., Neeteson, J.J., 2003. An evaluation of whole-farm nitrogen balances and related indices for efficient nitrogen use. Eur. J. Agron. 20,33–44.

Smith, V.H., Tilman, G.D., Nekola, J.C., 1999. Eutrophication: impacts of excess nutrient inputs on freshwater, marine, and terrestrial ecosystems. Environ. Pollut. 100, 179–196.

Spears, R.A., Kohn, R.A., Young, A.J., 2003. Whole-farm nitrogen balance on western dairy farms. J. Dairy Sci. 86, 4178–4186.

SPSS, 2015. Statistical Package for Social Sciences. SPSS Software Package 22.0 for Microsoft Windows, Chicago, Illinois, USA.

Suh, S. and Yee, S., 2011. Phosphorus use-efficiency of agriculture and food system in the US. Chemosphere. 84, 806-813.

Veeteelt,2015.<http://veeteelt.nl/nieuws/kringloopwijzer-2016-verplicht-voor-alle melkveebedrijven> Last accessed on 02-11-2015.

de Vries, W., Kros, J., Dolman, M.A., Vellinga, Th.V., de Boer, H.C., de Boer, I.J.M., Sonneveld, M.P.W., Bouma, J., 2015. Environmental impacts of innovative dairy farming

systems aiming at improved internal nutrient cycling: a multi-scale assessment. Sci. Total Environ. 536, 432-442.

Walker, W.E., Harremoës, P., Rotmans, J., van der Sluijs, J.P., van Asselt M.B.A., Janssen, P., Krayer von Krauss, M.P., 2003. Defining Uncertainty: A Conceptual Basis for Uncertainty Management in Model-Based Decision Support. Integr. Assess. 4, 5–17.

Whitehead, D.C., 1995. Volatilization of ammonia. In: Whitehead, D.C., ed. Grassland nitrogen. Wallingford, CAB International. 152-179.

Chapter 5

Assessing the impact of uncertainty on benchmarking the eco-efficiency of dairy farming using fuzzy data envelopment analysis

W. Mu[1,2], A.Kanellopoulos[2], C.E. van Middelaar[1], D. Stilmant[3], J.M. Bloemhof[2]

[1] Animal Production Systems group, Wageningen University, P.O. Box 338, 6700 AH Wageningen, the Netherlands

[2] Operations Research and Logistics group, Wageningen University, P.O. Box 8130, 6700 EW Wageningen, the Netherlands

[3] Farming Systems, Territory and Information Technologies Unit, Walloon Agricultural Research Centre, 100 rue du Serpont, 6800 Libramont, Belgium

This Chapter is submitted to Journal of Cleaner Production

Abstract

Data envelopment analysis (DEA) has been used to assess the eco-efficiency of dairy farming by taking into account multiple economic and environmental indicators. A limitation of standard DEA is that indicators are assumed to be deterministic and uncertainty is not taken into account. In practice, however, data used to calculate the economic and environmental performance of dairy farms, contain high levels of uncertainty. Neglecting uncertainty could cause biased benchmarking results and lead to wrong conclusions. Compared to the standard DEA, the advantage of fuzzy DEA is that it allows all parameters to be a range (i.e. minimum and maximum threshold values) instead of a single crisp number and by this we could allow uncertainties associated with the model parameters to be considered. The objective of this study is therefore to demonstrate how fuzzy DEA can be used to evaluate the eco-efficiency of dairy farming accounting explicitly for multiple indicators and corresponding uncertainty. In this study, we used fuzzy DEA to evaluate the eco-efficiency of 55 specialized dairy farms from different regions across Western Europe. We used N surplus, P surplus, land use, energy use as the environmental indicators and gross margin as the economic indicator. We first applied the standard DEA model on the aforementioned indicators to derive the eco-efficiency scores. Afterwards, we applied fuzzy DEA to explore the effect of uncertainty. We found that fuzzy DEA identified less efficient farms compared to the standard DEA. Fuzzy DEA provides different ranking results compared to the ranking results based on standard DEA. Same findings can be held when comparing other outputs of fuzzy DEA with the standard DEA . With fuzzy DEA, we found improvement percentage for each model input parameters are defined as ranges rather than a crisp value. In addition, fuzzy DEA identified different set peers and the importance of the peers are also different when compared to the peers of the standard DEA. All the aforementioned findings showed the importance of taking uncertainty into consideration in the benchmarking process, and how fuzzy DEA can be used to do so. With fuzzy DEA, we can identify the farms whose eco-efficiency performances are sensitive to the uncertainty of the inputs and outputs. Moreover, we demonstrated the type of outputs and managerial insights resulting from fuzzy DEA, and how they can be used by decision makers

to benchmark farm practices and improve the eco-efficiency of dairy production systems. The decision makers therefore should use the outcome of fuzzy DEA, which is more robust and reliable because uncertainty has been taken into account during the quantification processes.

5.1. Introduction

Global milk production has increased during the last decades, from 500 million tonnes in 1983 to 769 million tonnes in 2013 (FAO, 2016). With the ongoing urbanisation and improving living standards, the demand for dairy products is continuously expanding (Gerosa and Skoet, 2012). To meet the growing demand, world milk production is expected to grow further to 831 million tonnes in 2017 (FAO, 2017).

This increase of milk production can contribute to resource (e.g. land, energy) scarcity and various environmental problems such as climate change, acidification and eutrophication (Mu et al., 2016). As a result, the dairy sector is challenged to increase its eco-efficiency, which means, minimizing environmental impacts, while maintaining economic viability (Dolman et al., 2014). To meet this challenge, benchmarking individual farms based on their eco-efficiency can be valuable to identify best farm practices and to provide guidance for improving farm performance (Stokes et al., 2007).

To benchmark farms based on their eco-efficiency, a set of economic and environmental indicators must be selected and quantified. For measuring economic performance, indicators such as income, gross margin, and production cost are often used (van Middelaar et al., 2011; Iribarren et al., 2011; Lebacq et al., 2015). For measuring environmental performance, indicators like nitrogen and phosphorus surpluses, global warming potential, acidification potential, eutrophication potential, land use and energy use have been proposed (Mu et al., 2016). To prevent biased conclusions (Fraser and Cordina, 1999; Stokes et al., 2007; Iribarren et al., 2011) and to account for potential trade-offs between environmental problems, multiple indicators need to be taken into account when quantifying the eco-efficiency of dairy farming.

Data envelopment analysis (DEA) has been used to evaluate the eco-efficiency of agricultural systems accounting for multiple environmental and economic indicators simultaneously (Picazo-Tadeo et al., 2011, Iribarren et al., 2011). Eco-efficient practices are identified and used as a benchmark for current inefficient practices (Egilmez et al., 2016). A limitation of

standard DEA is that indicators are assumed to be deterministic and uncertainty is not taken into account (Stokes et al., 2007; Shortall et al., 2013). In practice, however, dairy systems involve complex production processes and some factors can be difficult to measure in a precise way (e.g. manure production, grass yield). Data used to calculate the economic and environmental performance of dairy farms, therefore, can contain high levels of uncertainty (Sefeedpari et al., 2012; Mugera, 2013). To avoid biased conclusions, it is important to take uncertainty into consideration when benchmarking the eco-efficiency of dairy farms (Mu et al., 2017). To the best of our knowledge, previous studies that used DEA to benchmark the eco-efficiency of dairy farms did not account for those uncertainties. As an alternative to standard DEA, Emrouznejad et al., (2014) suggest fuzzy DEA as a prominent tool to deal with the uncertainty of input and output data. Fuzzy DEA, therefore, could be a valuable approach to account for uncertainties when benchmarking the eco-efficiency of dairy farming.

The objectives of this study are to assess the impact of uncertainty on benchmarking the eco-efficiency of dairy farming, and to demonstrate how fuzzy DEA can be used to account for the uncertainty of multiple indicators. In section 5.2, theories on eco-efficiency, DEA and fuzzy DEA are introduced. Section 5.3 shows how fuzzy DEA can be used to quantify and benchmark the eco-efficiency of dairy farming while accounting for uncertainty based on a case study. The results of the case study are reported in section 5.4. In section 5.5, discussions on the important findings and conclusions of the study are provided.

5.2 Theory

5.2.1 Eco-efficiency of dairy farming based on a single environmental indicator

The concept of eco-efficiency was introduced in 1970s as the concept of "environmental efficiency" (Freeman et al., 1973). Later on, it was further developed to be the ratio between the economic value and environmental impact added by a system (Keffer and Shimp 1999; Seppälä et al., 2005; Zhang et al., 2008). Eco-efficiency has been widely quantified in the field of agriculture and food (Meul et al., 2007; van Middelaar et al., 2011; Müllera et al., 2015).

Net profit and gross value added are often used to quantify the economic dimension of eco-efficiency, while in case of the environmental dimension, indicators vary across studies. To quantify the environmental performance of dairy systems, the nutrient balance (NB) approach and life cycle assessment (LCA) are two commonly used methods (Oenema et al., 2003; de Boer, 2003; Basset-Mens et al., 2009). An NB quantifies the difference between nutrient flows entering and leaving a system, and is often used to determine losses of important nutrients, such as nitrogen (N) and phosphorus (P) (Mu et al., 2016). Based on an NB approach, indicators such as the nutrient use efficiency or the nutrient surplus (i.e., per hectare or per unit of product) of a system can be quantified. LCA is an internationally accepted and standardized method (ISO 14040, ISO 14041, ISO 14042, ISO 14043) that quantifies potential environmental impacts caused by the emission of pollutants and the use of resources during the entire life cycle of a product. Examples of LCA indicators are global warming potential, acidification potential, eutrophication potential, land use and energy use (Mu et al., 2016).

Generally, studies determine the eco-efficiency of a farm based on an individual environmental indicator (e.g. Basset-Mens et al., 2009; van Middelaar et al., 2011; Thanawong et al., 2014). Using one environmental indicator, however, might result in biased conclusions. For instance, a farm with low N surplus might have high energy use. Multiple environmental indicators are therefore needed to be taken into account when quantifying the eco-efficiency of dairy farming.

5.2.2 Quantifying eco-efficiency of dairy farming based on multiple environmental indicators using DEA

Data envelopment analysis (DEA) is a non-parametric approach which was originally developed by Charnes et al. (1978) to derive the efficiency score of a decision making unit (DMU), e.g. a dairy farm, by taking into consideration multiple indicators simultaneously.

To shortly explain the main principle of DEA, we use an illustrative two dimensional example, presented in Figure 5.1. We assume that a set of 5 dairy farms (A,B,C,D,E) are evaluated with respect to their N surplus and gross margin. Each farm is compared with all other farms based

on the gross margin produced per unit of N surplus. In this example, farm B has the highest gross margin per unit of N surplus and consequently it can be considered to be efficient. The line that crosses the origin as well as point B has a slope equal to the maximum amount of gross margin per unit of N surplus and is called the efficient frontier. In this example, we assume that the capacity of a farm to produce gross margin is independent of the production scale and the size of the operation. In other words, we assume a constant returns to scale (CRS), and the aforementioned frontier is the CRS eco-efficient frontier. All farms that are located on the frontier get an efficiency score of 1, while all other farms receive a non-negative efficiency score lower than 1, which is dependent on the distance to the efficient frontier. To increase its eco-efficiency, farm E, for example, has to move to point E_1 by increasing its gross margin (output orientation) without increasing its N surplus, or it has to move to point E_2 by decreasing its N surplus (input orientation) without decreasing its gross margin.
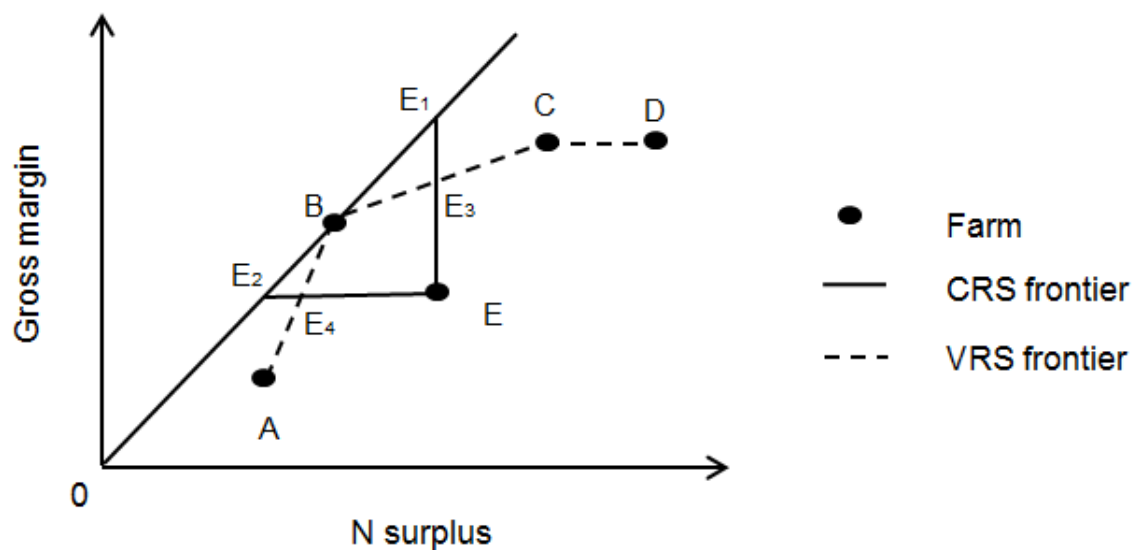
Figure 5.1 Illustrative example of Data Envelopment Analysis

In practice, the level of eco-efficiency of a DMU, in this case the gross margin per unit of N surplus, might depend on the scale of production (Spillman, 1923). The first units of fertilizer inputs, for example, might be more effective in producing outputs than the last ones. A variable returns to scale (VRS) DEA model takes into account the scale of production by assigning

efficiency scores based on a VRS frontier (Figure 5.1). Given the VRS efficient frontier of Figure 5.1, farms A, B and C are efficient. Farm D is weakly efficient which means that although it has an efficiency score of 1, it can still be improved by moving towards farm C. In the case of an output orientation VRS DEA model, farm E can increase its efficiency by moving to point E_3, using farm B and C as benchmarks. This means that management practices of farm E_3 will be based on a linear combination of the practices of farms B and C. In case of input orientation, farm E can increase its efficiency by moving to point E_4 using Farm A and B as benchmarks.

To generalize the aforementioned 2-dimensional example to a multi-dimensional example, we provide the mathematical formulation of the input oriented standard DEA model (1)-(6), which is used to evaluate the efficiency score of each farm with the efficiency score of all other farms in the sample.

$$Min\ \theta - \varepsilon(\sum_{i=1}^{m} s_i^- + \sum_{r=1}^{s} s_r^+) \tag{1}$$

$$s.t.\ \sum_{j=1}^{n} \lambda_j x_{ij} + s_i^- = \theta x_{i0} \quad i = 1, \dots, m \tag{2}$$

$$\sum_{j=1}^{n} y_{rj} \lambda_j - s_r^+ = y_{r0} \quad r = 1, \dots, s \tag{3}$$

$$\sum_{j=1}^{n} \lambda_j = 1 \qquad\qquad j = 1, \dots, n \tag{4}$$

$$\lambda_j \geq 0 \qquad\qquad j = 1, \dots, n \tag{5}$$

$$s_i^-, s_r^+ \geq 0 \tag{6}$$

Where θ is the efficiency score of the evaluated farm, ε is a very small positive number, s_i^- is the slack of input i and represents input excess, s_r^+ is the slack of output r and represents output shortfall. The slacks are used to avoid weak efficiencies. Assume that we have n DMUs noted as DMU_j (j=1,2,...,n), then x_{ij} (i=1,2,...,m) represents the ith input (i.e. less-is-better

indicator) for j^{th} DMU and y_{rj} *(r=1,2,...,s)* represent the rth output (i.e. more-is-better indicator) for j^{th} DMU. The fundamental principle of this DEA model is to let each DMU select the most favourable weights (i.e. λ_j) that can minimize the distance to the eco-efficiency frontier and derive the efficiency score θ. The constraint $\sum_{j=1}^{n} \lambda_j = 1$ imposes VRS, such that the scale of production is taken into account and farms are only compared to other farms of a similar size. Input orientation means inputs are minimized while the output is fixed. The standard DEA Model is solved with a two-phase method as described in Banker et al. (1984). We can apply this model to each DMU to derive its efficiency scores. A DMU is only efficient when $\theta = 1$ and $s_i^-, s_r^+ = 0$.

DEA has been used to assess the eco-efficiency of dairy farms by taking into account multiple economic and environmental indicators (De Koeijer et al., 2002; Ramilan et al., 2011). De Koeijer et al., (2002) applied DEA to evaluate the economic and environmental performance of arable farms in the Netherlands. Nitrogen surplus per ha and application of herbicides per ha were used as the main environmental indicators. Ramilan et al. (2011) assessed the economic and environmental efficiency of 210 dairy farms in New Zealand. In this study, nitrogen discharge per ha was used as environmental indicator. Results showed that farms that are economically efficient can still improve their environmental performance.

Several studies showed that standard DEA is very sensitive to (measurement) errors and that small changes in data can alter the efficiency results significantly (Stokes et al., 2007, Shortall etl., 2013). Because of the high uncertainty related to the data used to calculate the environmental performance of dairy systems fuzzy DEA can be used.

Fuzzy DEA has been proposed in industrial engineering, banking, and agricultural economics as an alternative variant of standard DEA, as it takes into account the uncertainty of model parameters (Srinivasa Raju and Nagesh Kumar, 2013; Emrouznejad et al., 2014; Egilmez et al., 2016; Wanke et al., 2016). So far, fuzzy DEA has not been used to evaluate the eco-efficiency of dairy systems.

5.2.3 Accounting uncertainty in benchmarking eco-efficiency of dairy farming by using fuzzy DEA

Accurate model inputs and outputs are vital for successful implementation of DEA. In practice, however, data can be imprecise which can lead to biased conclusions. A standard DEA model assumes that data are crisp, which means that data is assumed to be without error and uncertainties are not taken into account. Fuzzy DEA has been proposed as an alternative to standard DEA, when data is imprecise and vague. Figure 5.2 provides a two dimensional example to explain the main principle of fuzzy DEA.
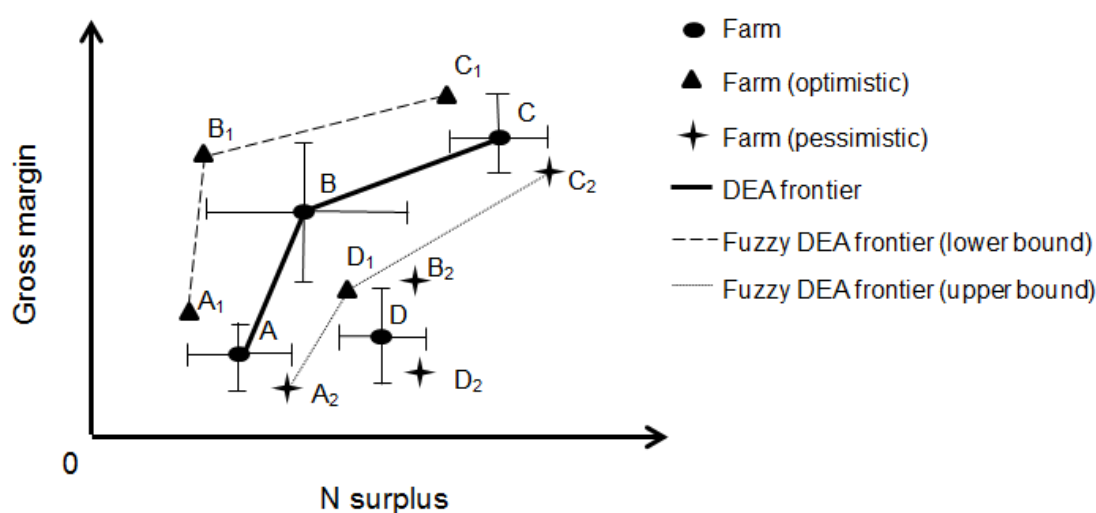
Figure 5.2 Fuzzy data envelopment analysis example

We assume that a set of four dairy farms is evaluated with respect to their N surplus and gross margin. Unlike the crisp (i.e. singe-valued) inputs and outputs of the standard DEA, in fuzzy DEA, inputs and outputs are ranges that are defined by the upper and lower bounds of the observed values (Figure 5.2). The dots represent the observed values of farm A, B, C and D, whereas the error bars represent the uncertainty range. The optimistic scenario, i.e., the scenario in which the farms have the highest efficiency, is based on the outcome representing the lowest inputs and highest outputs based on the uncertainty ranges. In Figure 5.2, A_1, B_1, C_1 and D_1 represent the optimistic scenario, while A_2, B_2, C_2 and D_2 represent the pessimistic scenario. Suppose farm D is the farm being evaluated. In this case, the lower bound of its

efficiency is represented by the situation in which farm D has the lowest efficiency (gross margin per kg of N surplus) (D_2) and the other farms have the highest efficiency (A_1, B_1, C_1). In this case, A_1, B_1 and C_1 define the new frontier, and farm D is not efficient. Unlike single-valued efficiency scores derived from standard DEA, fuzzy DEA provides a range of efficiency scores per farm.

The only difference in the general structure of the fuzzy DEA model compared to the standard DEA model is that inputs and outputs of the DMUs become fuzzy (i.e. $\tilde{x}_{ij}$ and $\tilde{y}_{rj}$ represent fuzzy inputs and outputs).

To solve the fuzzy DEA model, we followed the method proposed by Kao and Liu (2000). This method using α-cuts, is based on Zadeh's extension principle to transform the fuzzy DEA model to a series of standard crisp DEA models. The basic principle is to specify the fuzzy inputs and outputs in bounded ranges. Mugera (2013) pointed out that triangular fuzzy number are commonly used because they can be specified by the decision maker, so the inputs and outputs of DMUs can be defined as triangular fuzzy numbers. Let $\tilde{x}_{ij} = (x_{ij}^m, x_{ij}^l, x_{ij}^u)$ and $\tilde{y}_{rj} = (y_{rj}^m, y_{rj}^l, y_{rj}^u)$, So that fuzzy data $\tilde{x}_{ij}$ and $\tilde{y}_{ij}$ are specified with centre (x_{ij}^m, y_{rj}^m), which can be the measured data, and the left (x_{ij}^l, y_{rj}^l) and right (x_{ij}^u, y_{rj}^u) spread of the measured data.

To transfer the fuzzy numbers to crisp numbers, the α-cuts approach was applied. Using α-cuts, the fuzzy data (uncertainty range) can be represented by different level of confidence intervals. In this way, the α-cuts, representing the level of confidence, can form sets of movable boundaries (Kao and Liu, 2000). Figure 5.3 provides an example to illustrate the principle of fuzzy numbers and α-cuts. On the horizontal axis, we see that we have measured data x_{ij}^m with a bounded uncertainty range specified as $[x_{ij}^l, x_{rj}^u]$. The vertical axis indicates the level of α-cuts. α=1 indicates that there is no uncertainty, so the measured data is accurate. α=0 takes the whole fuzzy range into consideration. Uncertainty is increased by moving α from 1 to 0 (Figure 5.3).
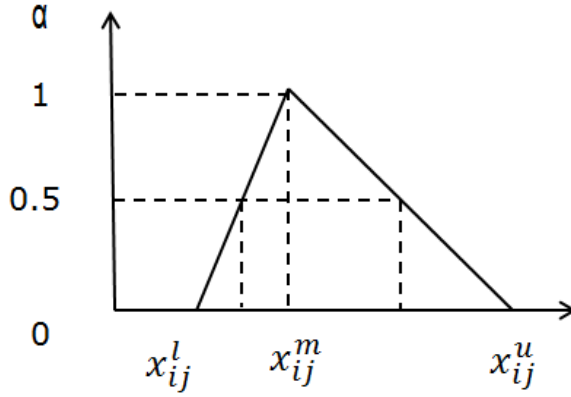
Figure 5.3 Illustration of α-cuts

At a given α-cut level ($0 \leqslant \alpha \leqslant 1$), the lower bound and upper bound of the fuzzy number $(\tilde{x}_{ij}, \tilde{y}_{rj})$ can be quantified as:

$$(x_{ij})_\alpha = \left[(x_{ij})^L_\alpha, (x_{ij})^U_\alpha\right] = \left[(x^m_{ij} - x^l_{ij})\alpha + x^l_{ij}, \ x^u_{ij} - (x^u_{ij} - x^m_{ij})\alpha\right] \tag{7}$$

$$(y_{rj})_\alpha = \left[(y_{rj})^L_\alpha, (y_{rj})^U_\alpha\right] = \left[(y^m_{rj} - y^l_{rj})\alpha + y^l_{rj}, \ y^u_{rj} - (y^u_{rj} - y^m_{rj})\alpha\right] \tag{8}$$

The lower and upper bound of the efficiency scores can be derived at a specified α level. According to Kao and Liu (2000), to find the lowest efficiency of DMU$_0$ compared to other DMUs, the output of DMU$_0$ and the input of all other DMUs should be set to their lowest values and the input of DMU$_0$ and the output of all other DMUs should be set to their highest values. This can be formulated mathematically as the fuzzy DEA model (9-14).

$$Min \ \theta - \varepsilon\left(\sum_{i=1}^{m} s^-_i + \sum_{r=1}^{s} s^+_r\right) \tag{9}$$

$$s.t. \ \sum_{\substack{j=1 \\ j \neq j_0}}^{n} \lambda_j x^L_{ij\alpha} + \lambda_{j_0} x^U_{i0\alpha} + s^-_i = \theta x^U_{i0\alpha} \quad i = 1, \dots, m \tag{10}$$

$$\sum_{\substack{j=1 \\ j \neq j_0}}^{n} \lambda_j y^U_{rj\alpha} + \lambda_{j_0} y^L_{r0\alpha} - s^+_r = y^L_{r0\alpha} \qquad r = 1, \dots, s \tag{11}$$

$$\sum_{j=1}^{n} \lambda_j = 1 \qquad\qquad\qquad j = 1, \dots, n \qquad\qquad\qquad (12)$$

$$\lambda_j \geq 0 \qquad\qquad\qquad j = 1, \dots, n \qquad\qquad\qquad (13)$$

$$s_i^-, s_r^+ \geq 0 \qquad\qquad\qquad\qquad\qquad (14)$$

The same principle can be applied to find the highest efficiency of DMU$_0$. By solving the fuzzy DEA model, we can derive an efficiency range at each level of α. The efficiency derived at $\alpha=1$ is the same as the one derived by standard DEA. The efficiency range derived at $\alpha=0$ means that the efficiency will never exceed the upper limit of the range or fall below the lower limit of the range.

5.2.4 Ranking method and eco-efficiency score

Fuzzy DEA produces efficiency scores as ranges at different α-cut levels. The scores might be difficult to interpret by decision makers, especially when the uncertainty level is unknown and too many *α* cuts are included. To facilitate the interpretation, the ranges at different α-cut levels need to be composed into crisp values. Chen and Klein, (1997) provide a solution to derive ranking indexes (I_j) by using equation (15):

$$I_j = \frac{\sum_{k=0}^{n}((\theta_j)_{\alpha_k}^{U} - c)}{\left[\sum_{j=0}^{n}\left((\theta_j)_{\alpha_k}^{U} - c\right) - \sum_{j=0}^{n}((\theta_j)_{\alpha_k}^{L} - d) \right]}, \quad n \to \infty \qquad\qquad (15)$$

Where θ_j is the efficiency score for DMU *j*, *U* and *L* represent the upper and lower bound of the efficiency ranges respectively; $c = \min\limits_{j,k}\left\{((\theta_{jk})_{\alpha_k}^{L})\right\}$, which is the lowest efficiency score among the lower bounds of the efficiency ranges at all α cuts, $d = \max\limits_{j,k}\left\{((\theta_{jk})_{\alpha_k}^{U})\right\}$, which is the largest efficiency score among the upper bounds of the efficiency ranges at all *α* cuts; and *n* is the number of *α* cuts.

5.3 Case study

5.3.1 Data description

In this study, we used fuzzy DEA to evaluate the eco-efficiency of 55 specialized dairy farms from different regions across Western Europe. Table 5.1 provides the description of the farm characteristics.

Table 5.1 Descriptive statistics of important farm characteristics

Characteristics	Unit	Mean	SD	Minimum	Maximum
Farm area	ha	73.6	36.8	35.8	264.8
Dairy cows	number	103.3	33.3	49	194
Milk production	kg cow^{-1}year^{-1}	7993.3	1144.6	5700	10742
Milk fat content	%	4.2	0.2	3.7	4.8
Milk protein content	%	3.5	0.1	3.1	3.7
Concentrate fed	kg cow^{-1}year^{-1}	1434.8	726.2	317	3445

We used total gross margin as the economic indicator, defined as revenue minus total variable cost. Environmental indicators have been calculated based on a life cycle assessment approach, including all processes from cradle to farm gate, and using farm data from 2010 (Dairyman, 2010; Mu et al., 2017). A detailed description of the calculation methods can be found in Mu et al. (2017). Environmental indicators included were N surplus, P surplus, land use, energy use, all expressed per kg of fat-and-protein corrected milk (FPCM) and based on economic allocation. Descriptive statistics of the economic and environmental indicators are presented in Table 5.2.

Table 5.2 Descriptive statistics of the economic and environmental indicators

Indicator[1]	Unit	Mean	SD	Minimum	Maximum
Gross margin	euro × kg FPCM^{-1}	0.2	0.1	0.1	0.4
N surplus	g N × kg FPCM^{-1}	14.4	5.2	4.9	27.0
P surplus	g P × kg FPCM^{-1}	1.3	0.9	0.1	4.9
Land use	m^2 × kg FPCM^{-1}	1.3	0.4	0.7	2.8
Energy use	MJ × kg FPCM^{-1}	2.8	0.8	1.8	5.1

[1] The system boundary for all environmental indicators is from cradle-to-farm gate

5.3.2 Set up of calculations

According to Pérez Urdiales et al. (2016), reducing environmental impacts while maintaining economic return is of major interest to farmers and policymakers because the profit margins as well as environmental policies become tighter over time. We therefore applied input oriented DEA models.

The environmental indicators are defined as "less is better" indicators and the economic indicator is defined as "more is better" indicator. First, we applied the VRS DEA model on the set of indicators as mentioned in Table 5.2 to derive the eco-efficiency score for each farm. Second, we applied the fuzzy DEA model to explore the effect of uncertainty. To determine the fuzzy lower and upper bound of the range of each indicator we assumed -20% and +20% of the observed level respectively. Preferably, the upper and lower bound are based on actual uncertainty values, but data requirements are high and limit the inclusion of such values. We used α-cuts to convert the fuzzy DEA model into the corresponding crisp model. We took into account the whole fuzzy range assuming 11 α-cuts with values that range from 0 to 1 and with pre-specified intervals of 0.1 (i.e. α=1, 0.9, 0.8,...,0.1,0), so the uncertainty range of each indicator can vary from 0-20%. By solving the crisp model, we derived a lower and upper bound of the efficiency score at each α-cut level. Third, two scenarios were defined to explore different levels of uncertainty by using different combinations of the α-cuts. In the first scenario, we

assumed low uncertainty and we included α-cut levels of 0.9, 0.8, and 0.7. In the second scenario, we assumed high uncertainty and we included α-cut levels of 0.1, 0.2, and 0.3. Then we used the method proposed by Chen and Klein (1997) to aggregate the efficiency scores at each α-cut into one single value (i.e. CK index) for each farm. The CK index can be used as the eco-efficiency score for benchmarking farms.

The ranking of farms based on the eco-efficiency scores for each of the two scenarios was compared with the ranking based on standard DEA. Subsequently, the improvement potential per farm based on standard DEA was compared with the improvement potential based on fuzzy DEA. In addition, standard DEA and fuzzy DEA were compared with regard to the number of times a farm was selected as a peer (i.e., a benchmark). The importance of peers were calculated by considering the frequency of a farm to be a peer for inefficient farms and the weight of the peer (i.e. the corresponding λi variables of model 10-12).

All the computations were made in Excel 2010, Xpress IVE and IBM SPSS statistics 22.

5.4 Results

At an α-cut level of 1 (i.e. equivalent to standard DEA), nine eco-efficient farms were identified. The higher the uncertainty level, the less farms were qualified as being eco-efficient. At α=0.9, six farms had an eco-efficiency score of 1 at both the lower and upper bound, at α=0.8, four farms had this score, and at α=0.7 three farms had this score. The same three farms kept this score at both the lower and upper bound until α=0.4. Only one farm had an efficiency score of 1 at both the lower and upper bound across all α-cuts levels, which means that this farm is most efficient in all situation, regardless the uncertainty. An overview of the fuzzy DEA results with 11 α cuts can be found in the Appendix.

Figure 5.4 shows how the uncertainty of model parameters can affect the eco-efficiency scores. The left and right figures show the efficiency scores derived from fuzzy DEA at α=0.9 and α=0 respectively. Farms that do not have an efficiency score of 1, have an efficiency range instead of one single score. For low levels of uncertainty (i.e. α=0.9), it's still possible to

compare farms and identify differences in eco-efficiency scores. However, in high uncertainty levels (α=0), it becomes more difficult to identify differences in the eco-efficiency scores among the farms. Results show that ignoring uncertainty can lead to biased conclusions when benchmarking farms based on their eco-efficiency, i.e. farms that are identified as efficient might not be efficient and vice versa, when results are uncertain because of measurement errors or temporal variability.
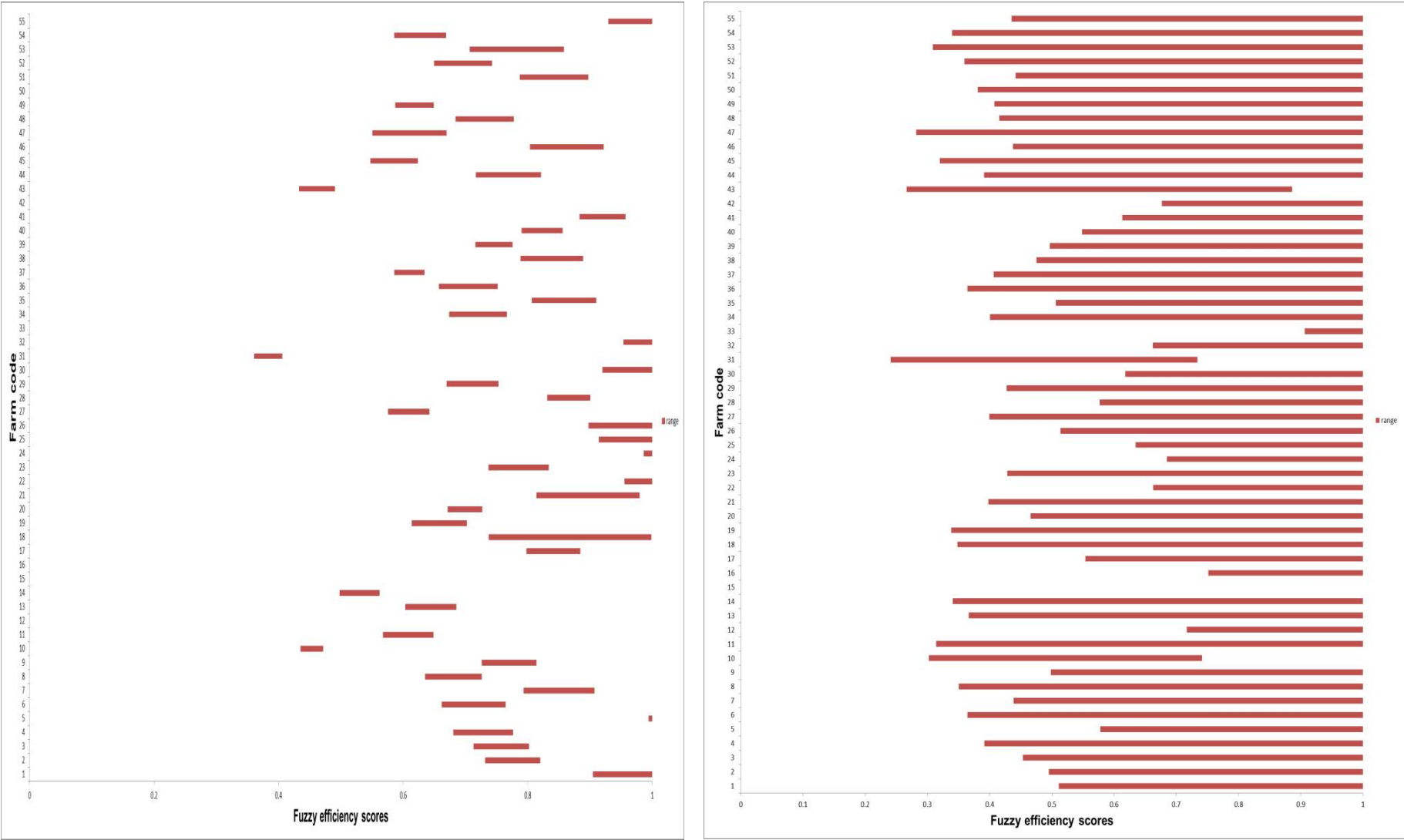
Figure 5.4 Eco-efficiency scores derived with fuzzy DEA at α=0.9 (left figure) and α=0 (right figure)

Figure 5.5 shows the ranking results of the low and high uncertainty scenarios based on fuzzy DEA compared to standard DEA. The farms located on the 1-1 line have the same ranking based on their eco-efficiency score with both fuzzy and standard DEA. The more the farm deviates from the 1-1 line, the more different the ranking based on the eco-efficiency score is. Compared to standard DEA which identified 9 eco-efficient farms, fuzzy DEA identified 3 (scenario 1) or 1 (scenario 2) eco-efficient farms. Figure 5.5 illustrates that compared to standard DEA, fuzzy DEA does not change the ranking results when uncertainty levels are low (scenario 1), but at higher levels of uncertainty (scenario 2), ranking results can change considerably.
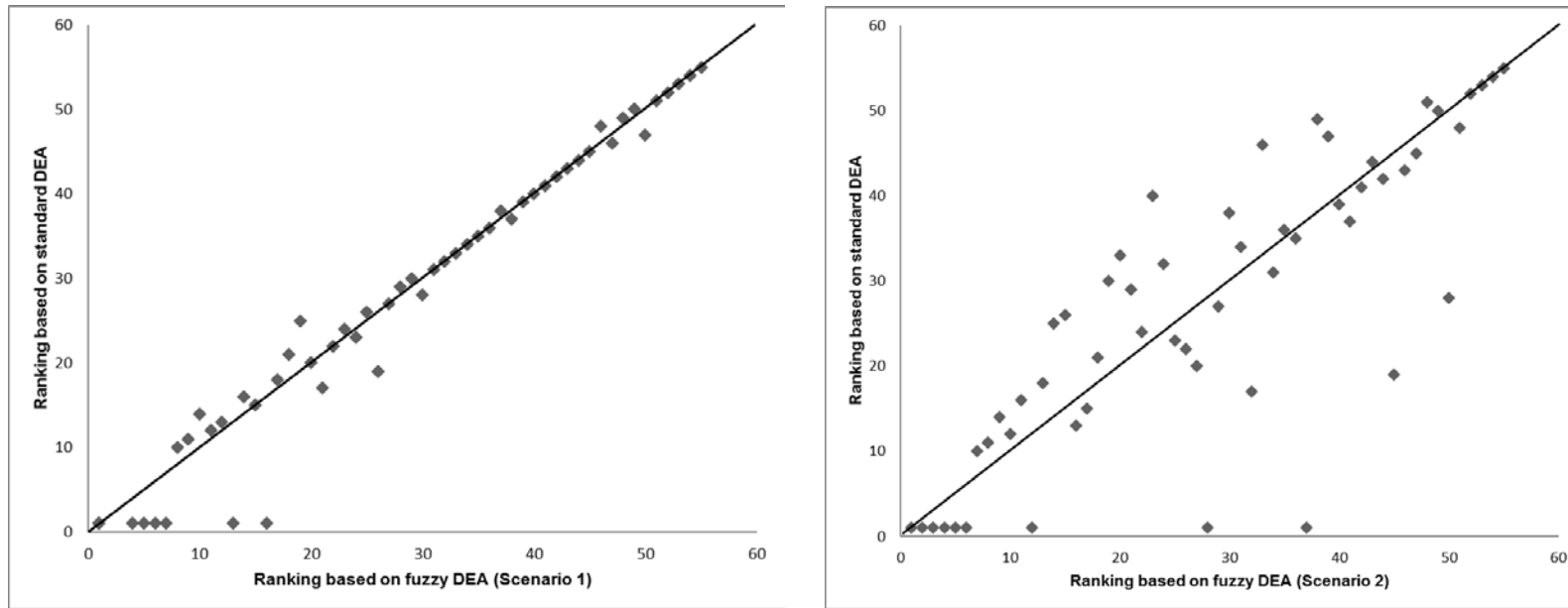
Figure 5.5 Comparison of the ranking of farms based on their eco-efficiency score by using standard DEA or fuzzy DEA (i.e. CK index) for scenario 1 (low uncertainty) and scenario 2 (high uncertainty).
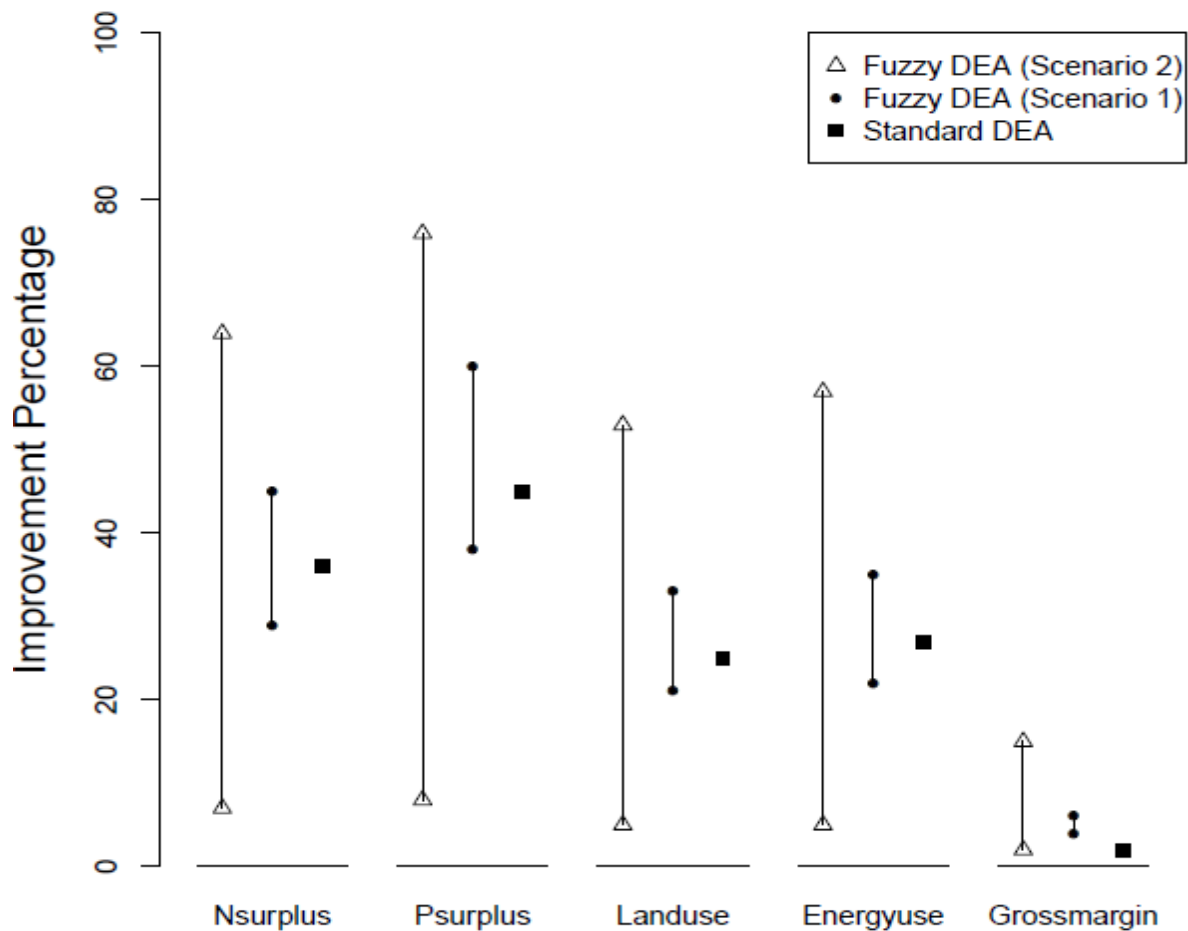
Figure 5.6 Total scope for improvement (%) of the 55 farms for different economic and environmental indicators calculated with standard DEA, and fuzzy DEA based on scenario 1 (low uncertainty) and scenario 2 (high uncertainty).

Figure 5.6 shows the total scope for improvement (%) if all 55 farms move to an efficiency score of 1. Results are expressed per environmental indicator for standard DEA and fuzzy DEA based on a low and a high level of uncertainty. Based on standard DEA, the improvement potential is represented by a single (crisp) value; based on fuzzy DEA, the improvement potential is represented by a range, which is defined by the lower and upper bound of the fuzzy efficiency frontier. Results of the high uncertainty scenario show that the actual improvement potential might be very different from the improvement potential calculated with standard DEA, in case indicators contain uncertainty.

Table 5.3 lists the identified peers and the importance of these peers for standard DEA and fuzzy DEA based on a low and a high level of uncertainty. Standard DEA resulted in a different set of peers than fuzzy DEA. The most important peers (i.e. farm 5, 12, 15, 33, 42) are the same for both standard and fuzzy DEA, but the relative importance of the peers varied between the two methods. The importance of peers (%) is calculated based on the frequency of a farm being a peer, and the weight of the peer (i.e. the corresponding λi variables of model 10-12). For example, the importance of peer farm 12 is higher based on standard DEA (10.7%) than based on fuzzy DEA (low uncertianty 9.0%, high uncertianty 10.3%). Correct identification of peers and their importance is relevant because the practices of peers are of first importance to improve the practices of other (less eco-efficient) farms. Based on detailed farm information, we can explore why certain farms are popular peers and provide specific managerial insights that will improve the overall eco-efficiency of the region.

Table 5.3 Importance of the peer farms under standard DEA and fuzzy DEA on scenario 1 (low uncertainty) and scenario 2 (high uncertainty).

Peer farm	Standard DEA	Fuzzy DEA (scenario 1)	Fuzzy DEA (scenario 2)
P(5)	9.4%	9.3%	5.6%
P(12)	10.7%	9.0%	10.3%
P(15)	12.1%	12.6%	10.6%
P(16)	-	0.8%	0.8%
P(22)	-	0.9%	0.9%
P(24)	3.5%	2.7%	1.5%
P(26)	-	-	0.4%
P(28)	-	-	0.5%
P(33)	45.4%	45.4%	56.1%
P(42)	15.7%	17.0%	12.7%
P(46)	-	-	0.3%
P(50)	0.5%	0.5%	0.2%
P(55)	2.9%	1.9%	-

Farm 33, for example, is the peer farm with the highest weight. This farm is a pasture based Irish dairy farm with a low input of fertilizer and feed. Milk production levels are relatively high (i.e., 7910 kg per cow per year), resulting a relatively high gross margin. Compared to concentrate-based farms, this farm has a low level of farm inputs. Compared to other pasture-based farms, the amount of concentrates per cow is relatively high, and so is the level of milk production per cow. Farm 33 can serve as an example for other farms that aim to increase their eco-efficiency.

5.5 Discussion and Conclusions

We explored the potential of fuzzy DEA to derive eco-efficiency scores and benchmark dairy farms. DEA-based benchmarking methods enable assessment of multiple economic and environmental indicators simultaneously. Dairy production involves complex production processes, and some indicators are difficult to measure, such as manure production or grass yields. This can cause imprecise estimations of the input and output parameters, which could lead to biased results. Therefore, accounting for uncertainty is important. Compared to standard DEA, which is deterministic and does not enable to account for uncertainty, fuzzy DEA has the advantage that it allows uncertainty to be included by expressing all parameters as ranges instead of single crisp numbers. Our findings are in line with literature that applied fuzzy DEA to account for data uncertainty in efficiency evaluation problems. In general, fuzzy DEA results in a lower number of efficient farms compared to standard DEA (Kao and Liu, 2005; Mugera, 2013; Egilmez et al., 2016). Based on a 20% uncertainty range, fuzzy DEA provided a different ranking of farms based on their eco-efficiency performance compared to the ranking of those farms based on standard DEA. In addition, the improvement percentage for each model input parameters is defined as a range rather than a crisp value when using fuzzy instead of standard DEA. Finally, fuzzy DEA resulted in a different set of peers compared to the peers identified with standard DEA. All the aforementioned findings show that decision makers should be aware that standard DEA can result in misleading and inaccurate conclusions when data used to calculate the indicators contain uncertainty (Sefeedpari et al., 2012). By taking uncertainty into account during the quantification processes, fuzzy DEA can contribute to increasing the reliable of results and prevent biased conclusions (Mugera, 2013; Egilmez et al., 2016).

In this study we focussed mainly on demonstrating the importance of accounting for uncertainty when benchmarking farms based on their eco-efficiency by means of DEA. We demonstrated how fuzzy DEA can be used to do so, and provided insight in the type of outputs and how they can be used for benchmarking dairy farms. We used farms from different regions with different bio-physical and social-economic conditions. This implies that given our dataset not all farms

are fully comparable and it is not always possible to identify farm specific improvement options based on farm comparison. However, the purpose of the study was not to conduct in-depth farm analysis and provide farm-specific advices after farm comparison, but rather to provide a numerical example to demonstrate the importance of taking uncertainty into consideration in the process of benchmarking, and to show what kind of results can be produced with fuzzy DEA. To identify farm specific improvement options, a more homogenous set of farms is required.

In this study we assumed an uncertainty range of 20% for all indicators and for all farms. With different α-cut scenarios, we were able to perform the sensitivity analysis within the 20% fuzzy ranges. To improve interpretation of results, sample specific uncertainty ranges of all the input and output data are required.

This study demonstrated how fuzzy DEA can be used to evaluate the eco-efficiency of dairy farming, accounting explicitly for multiple indicators and corresponding uncertainty. We showed the importance of taking uncertainty into consideration in the benchmarking process, and how fuzzy DEA can be used to do so. With fuzzy DEA, we can identify the farms whose eco-efficiency performances are sensitive to the uncertainty of the inputs and outputs. Moreover, we demonstrated the type of outputs and managerial insights resulting from fuzzy DEA, and how they can be used by decision makers to benchmark farm practices and improve the eco-efficiency of dairy production systems.

Appendix

Table. An overview of the fuzzy DEA results with 11 α cuts

Farm code	α=1		α=0.9		α=0.8		α=0.7		α=0.6		α=0.5		α=0.4		α=0.3		α=0.2		α=0.1		α=0	
	L	U	L	U	L	U	L	U	L	U	L	U	L	U	L	U	L	U	L	U	L	U
BF05	0.965128	0.965128	0.90487	1	0.848655	1	0.796158	1	0.747085	1	0.701171	1	0.658174	1	0.617874	1	0.580072	1	0.544585	1	0.511246	1
BF06	0.774583	0.774583	0.731607	0.820406	0.691253	0.869323	0.658248	0.921974	0.632312	0.979403	0.60732	1	0.58322	1	0.559965	1	0.537513	1	0.515822	1	0.494853	1
BF07	0.756353	0.756353	0.713417	0.802295	0.673154	0.853388	0.635357	0.910264	0.599836	0.971453	0.566421	1	0.534956	1	0.512789	1	0.492228	1	0.472364	1	0.453162	1
BF08	0.726927	0.726927	0.680593	0.776698	0.637414	0.83022	0.597134	0.889441	0.560372	0.95578	0.527095	1	0.496087	1	0.467454	1	0.440507	1	0.415124	1	0.391195	1
BF09	1	1	0.994346	1	0.931774	1	0.87338	1	0.82191	1	0.774857	1	0.730619	1	0.688991	1	0.649783	1	0.612823	1	0.577954	1
BF11	0.711545	0.711545	0.662263	0.764692	0.616522	0.822064	0.57403	0.884062	0.535756	0.951132	0.501211	1	0.468943	1	0.439627	1	0.412866	1	0.387737	1	0.364122	1
BF12	0.848311	0.848311	0.793466	0.907264	0.742393	0.970703	0.694784	1	0.650363	1	0.608881	1	0.57011	1	0.533844	1	0.499895	1	0.468093	1	0.43828	1
BF13	0.679229	0.679229	0.635191	0.726572	0.594187	0.777525	0.55597	0.862298	0.520318	0.988863	0.48703	1	0.455921	1	0.426828	1	0.399619	1	0.374258	1	0.35048	1
BW10	0.768945	0.768945	0.726407	0.814294	0.690214	0.863355	0.663083	0.917021	0.636957	0.974558	0.611781	1	0.587504	1	0.564078	1	0.541461	1	0.51961	1	0.498488	1
BW11	0.453157	0.453157	0.435386	0.471653	0.418299	0.49092	0.401856	0.511007	0.386022	0.531967	0.370765	0.553858	0.356052	0.577141	0.341855	0.611981	0.328148	0.651687	0.314906	0.694672	0.302105	0.741159
BW12	0.606906	0.606906	0.567759	0.648982	0.531299	0.694254	0.497308	0.743021	0.46559	0.795614	0.435966	0.852405	0.408275	0.913811	0.38237	0.982939	0.358116	1	0.335393	1	0.314088	1
BW14	1	1	1	1	0.992168	1	0.953168	1	0.915612	1	0.879422	1	0.844524	1	0.810851	1	0.778339	1	0.746929	1	0.716566	1
BW16	0.643043	0.643043	0.603431	0.685521	0.56645	0.731122	0.53189	0.780135	0.499561	0.832879	0.46929	0.889713	0.440921	0.951037	0.415391	1	0.398074	1	0.38201	1	0.366481	1
BW17	0.529026	0.529026	0.497866	0.562366	0.47147	0.598074	0.452938	0.636646	0.435092	0.678319	0.417894	0.723154	0.401311	0.771457	0.38531	0.823572	0.36986	0.879887	0.354935	0.940836	0.340506	1
BW18	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
FB01	1	1	1	1	1	1	0.999251	1	0.959879	1	0.921939	1	0.885355	1	0.850053	1	0.815969	1	0.783041	1	0.75121	1
FB10	0.834249	0.834249	0.798283	0.884795	0.766953	0.938822	0.736806	0.997841	0.707775	1	0.6798	1	0.652823	1	0.626794	1	0.601662	1	0.577382	1	0.553911	1
GE06	0.861693	0.861693	0.737414	0.998891	0.62488	1	0.553944	1	0.51823	1	0.484972	1	0.453897	1	0.424839	1	0.397647	1	0.372183	1	0.348321	1
GE09	0.656529	0.656529	0.613873	0.702391	0.574161	0.751754	0.537151	0.805377	0.502629	0.866262	0.4704	0.932195	0.440285	1	0.412123	1	0.385769	1	0.361088	1	0.337958	1
GE10	0.698814	0.698814	0.67141	0.727337	0.645059	0.765363	0.619703	0.810313	0.595286	0.85836	0.571757	0.909789	0.549068	0.964917	0.527176	1	0.506038	1	0.485617	1	0.465876	1
GE11	0.893599	0.893599	0.814334	0.979933	0.741522	1	0.674609	1	0.619872	1	0.575944	1	0.535099	1	0.497097	1	0.461719	1	0.428765	1	0.398055	1
IR01	0.994568	0.994568	0.955565	1	0.918062	1	0.881975	1	0.847224	1	0.813737	1	0.781446	1	0.750288	1	0.720204	1	0.69114	1	0.663045	1
IR02	0.783967	0.783967	0.737229	0.834	0.69352	0.887621	0.652599	0.945158	0.614251	1	0.578278	1	0.544501	1	0.512758	1	0.4829	1	0.454792	1	0.42831	1
IR03	1	1	0.986687	1	0.947963	1	0.9107	1	0.874817	1	0.84024	1	0.806897	1	0.774724	1	0.74366	1	0.71365	1	0.68464	1
IR04	0.961904	0.961904	0.914281	1	0.878399	1	0.84387	1	0.810621	1	0.778581	1	0.747685	1	0.717873	1	0.689089	1	0.66128	1	0.634399	1
IR05	0.958274	0.958274	0.89777	1	0.841358	1	0.788708	1	0.739522	1	0.693651	1	0.650845	1	0.610738	1	0.573129	1	0.539245	1	0.514033	1
IR06	0.60704	0.60704	0.575839	0.642367	0.553239	0.680046	0.531492	0.720284	0.510551	0.763313	0.490371	0.809388	0.470912	0.858797	0.452135	0.91186	0.434007	0.969636	0.416492	1	0.399562	1
IR07	0.865341	0.865341	0.831406	0.900661	0.798776	0.937452	0.767378	0.97581	0.737142	1	0.708006	1	0.679911	1	0.652801	1	0.626626	1	0.601338	1	0.576894	1
IR08	0.710314	0.710314	0.670151	0.753183	0.63248	0.803884	0.597108	0.860746	0.563858	0.92209	0.532573	0.98835	0.503789	1	0.483444	1	0.464059	1	0.445332	1	0.427229	1
IR09	0.976048	0.976048	0.920255	1	0.868214	1	0.821653	1	0.789279	1	0.758082	1	0.728	1	0.698972	1	0.670946	1	0.64387	1	0.617697	1
IR10	0.382588	0.382588	0.360684	0.405984	0.340279	0.432038	0.321474	0.460576	0.307404	0.491265	0.295253	0.524311	0.283537	0.559943	0.272232	0.598419	0.261316	0.64003	0.250771	0.685102	0.240577	0.734005
IR11	0.992993	0.992993	0.954052	1	0.916609	1	0.880579	1	0.845883	1	0.812449	1	0.780209	1	0.7491	1	0.719064	1	0.690046	1	0.661996	1
IR12	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	0.984389	1	0.944664	1	0.906263	1
IR13	0.718898	0.718898	0.674046	0.767025	0.632201	0.818725	0.593123	0.874326	0.556592	0.934198	0.522411	0.99875	0.4904	1	0.462116	1	0.435804	1	0.417589	1	0.400614	1
IR14	0.856907	0.856907	0.806781	0.910511	0.759855	0.967903	0.715877	1	0.674619	1	0.635874	1	0.599473	1	0.573047	1	0.55007	1	0.527872	1	0.506414	1
IR15	0.703276	0.703276	0.657765	0.752199	0.615384	0.804848	0.575881	0.861568	0.539025	0.922749	0.504609	1	0.472444	1	0.442358	1	0.414197	1	0.387817	1	0.363993	1
IR16	0.609736	0.609736	0.585825	0.634623	0.562833	0.660548	0.540709	0.687575	0.519405	0.715777	0.498875	0.751361	0.479078	0.799977	0.459976	0.852674	0.441533	0.909548	0.423715	0.973459	0.406491	1
IR17	0.837413	0.837413	0.788865	0.889305	0.743394	0.944835	0.700757	1	0.660738	1	0.623137	1	0.588754	1	0.557609	1	0.528142	1	0.500239	1	0.475284	1
IR18	0.745358	0.745358	0.716128	0.77578	0.688022	0.807471	0.660978	0.84051	0.634934	0.874985	0.609838	0.910993	0.585638	0.948637	0.562287	0.988032	0.539742	1	0.51796	1	0.496905	1
IR20	0.82262	0.82262	0.790361	0.856197	0.759342	0.891172	0.729493	0.927636	0.700751	0.965685	0.673053	1	0.646344	1	0.620573	1	0.59569	1	0.571651	1	0.548413	1
LU02	0.919592	0.919592	0.883529	0.957126	0.848854	0.996224	0.815487	1	0.783356	1	0.752393	1	0.722536	1	0.693727	1	0.665911	1	0.639038	1	0.613061	1
NL01	1	1	1	1	1	1	1	1	1	1	1	1	0.946187	1	0.87094	1	0.801244	1	0.736671	1	0.676827	1
NL03	0.460778	0.460778	0.432911	0.490632	0.40687	0.522651	0.38251	0.557033	0.359699	0.594	0.338319	0.633798	0.318852	0.676702	0.301799	0.723022	0.289698	0.773107	0.278008	0.827347	0.266706	0.886187
NL04	0.767313	0.767313	0.71679	0.821671	0.669782	0.880214	0.626005	0.943337	0.585198	1	0.547128	1	0.511583	1	0.478368	1	0.447309	1	0.418245	1	0.39103	1
NL05	0.584388	0.584388	0.547606	0.623873	0.513306	0.666308	0.481288	0.711965	0.452842	0.761148	0.426781	0.814199	0.402287	0.871497	0.379245	0.933471	0.357549	1	0.337653	1	0.319679	1
NL06	0.86091	0.86091	0.803976	0.922177	0.751015	0.991741	0.701705	1	0.655751	1	0.612889	1	0.572878	1	0.535501	1	0.500558	1	0.467868	1	0.437267	1
NL07	0.607878	0.607878	0.550931	0.67002	0.501424	0.737864	0.461865	0.832318	0.429194	0.964802	0.399183	1	0.372429	1	0.347468	1	0.324166	1	0.302396	1	0.282047	1
NL08	0.729394	0.729394	0.684232	0.777835	0.642082	0.829852	0.602702	0.885773	0.565875	0.945967	0.532745	1	0.501743	1	0.473333	1	0.451113	1	0.432909	1	0.415311	1
NL09	0.613511	0.613511	0.587464	0.649211	0.564408	0.687633	0.542223	0.730939	0.520858	0.777398	0.500271	0.827309	0.480419	0.881003	0.461264	0.938852	0.442769	1	0.424901	1	0.407628	1
NL10	1	1	1	1	0.764382	1	0.694063	1	0.629463	1	0.570104	1	0.515714	1	0.472363	1	0.438719	1	0.40868	1	0.380629	1
NL11	0.840579	0.840579	0.787495	0.897573	0.738002	0.958834	0.691811	1	0.648659	1	0.60831	1	0.570549	1	0.535179	1	0.502024	1	0.470921	1	0.441722	1
NL12	0.69467	0.69467	0.649812	0.742886	0.608035	0.794769	0.56909	0.850658	0.532751	0.910937	0.498813	0.97603	0.467092	1	0.437418	1	0.409638	1	0.383612	1	0.359213	1
NL13	0.779101	0.779101	0.707228	0.858559	0.641299	0.996589	0.580803	1	0.525277	1	0.474303	1	0.428184	1	0.388184	1	0.358343	1	0.332683	1	0.30878	1
NL15	0.625976	0.625976	0.585877	0.66906	0.548517	0.715402	0.515691	0.765304	0.485584	0.819105	0.457331	0.877183	0.430794	0.93996	0.405845	1	0.382368	1	0.360259	1	0.339421	1
NL16	1	1	0.929768	1	0.844024	1	0.765314	1	0.693041	1	0.628021	1	0.583235	1	0.541577	1	0.502806	1	0.466702	1	0.43502	1

5.6 References

Banker, R.D., Charnes, A., Cooper, W.W., 1984. Some models for estimating technical and scale inefficiencies in data envelopment analysis. Manag. Sci. 30, 1078-1092.

Basset-Mens, C., Ledgard, S., Boyes, M., 2009. Eco-efficiency of intensification scenarios for milk production in New Zealand. Ecol. Econ. 68, 1615-1625.

Charnes, A., Cooper, W.W., Rhodes, E., 1978. Measuring the efficiency of decision making units. EUR. J. OPER. RES. 2, 429-444.

Chen, C.-B., Klein, C.M., 1997. A simple approach to ranking a group of aggregated fuzzy utilities. IEEE. Trans. Syst. Man. Cybern. B. Cybern. 27, 26-35.

Ciriacy-Wantrup, S., 1971. The economics of environmental policy. Land Economics 47, 36-45.

De Boer, I.J., 2003. Environmental impact assessment of conventional and organic milk production. Livest. Prod. Sci. 80, 69-77.

De Koeijer, T., Wossink, G., Struik, P., Renkema, J., 2002. Measuring agricultural sustainability in terms of efficiency: the case of Dutch sugar beet growers. J. Environ. Manage. 66, 9-17.

Dolman, M., Sonneveld, M., Mollenhorst, H., De Boer, I., 2014. Benchmarking the economic, environmental and societal performance of Dutch dairy farms aiming at internal recycling of nutrients. J. Clean. Prod. 73, 245-252.

Emrouznejad, A., Tavana, M., Hatami-Marbini, A., 2014. Performance measurement with fuzzy data envelopment analysis. Studies in Fuzziness and Soft Computing, Springer-Verlag, Berlin Heidelberg.

Egilmez, G., Gumus, S., Kucukvar, M., Tatari, O., 2016. A fuzzy data envelopment analysis framework for dealing with uncertainty impacts of input–output life cycle assessment models on eco-efficiency assessment. J. Clean. Prod. 129, 622-636.

FAO, 2016. Food outlook biannual report on global food markets. http://www.fao.org/3/a-I5703E.pdf.

FAO, 2017. Food outlook biannual report on global food markets. http://www.fao.org/3/a-i7343e.pdf.

Freeman, M.A., Haveman, K., Kneese, A.V., 1973. The Economics of Environmental Policy. John Wiley and Sons, New York.

Fraser, I., Cordina, D., 1999. An application of data envelopment analysis to irrigated dairy farms in Northern Victoria, Australia. Agric. Sys. 59, 267-282.

Gerosa, S., Skoet, J., 2012. Milk availability. Trends in production and demand and medium-term outlook. FAO, Agricultural Development Economics Div. http://www.fao.org/3/a-an450e.pdf.

Iribarren, D., Hospido, A., Moreira, M.T., Feijoo, G., 2011. Benchmarking environmental and operational parameters through eco-efficiency criteria for dairy farms. Sci. Total Environ. 409, 1786-1798.

ISO, I., 1997. 14040: Environmental Management: Life Cycle Assessment: Principles and Framework. European Committee for Standardization (CEN), Brussels, Belgium (1997).

ISO, I., 1998. 14041: Environmental management-life cycle assessment-goal and scope definition and inventory analysis. European Committee for Standardization (CEN), Brussels, Belgium (1998).

ISO, I., 2000. 14042: environmental management—life cycle assessment—life cycle impact assessment. Geneva, Switzerland. European Committee for Standardization (CEN), Brussels, Belgium (2000).

ISO, I., 2000. 14043: Environmental Management–Life Cycle Assessment–Life Cycle Interpretation. European Committee for Standardization (CEN), Brussels, Belgium (2000).

Kao, C., Liu, S.-T., 2000. Fuzzy efficiency measures in data envelopment analysis. Fuzzy. Set. Syst. 113, 427-437.

Kao, C., Liu, S.-T., 2005. Data envelopment analysis with imprecise data: an application of Taiwan machinery firms. Int. J. Uncertainty Fuzziness Knowledge Based Syst. 13, 225-240.

Keffer, C., Shimp, R., Lehni, M., 1999. Eco-efficiency indicators and reporting. World Business Council for Sustainable Development (WBCSD), London.

Karsak, E.E., 2008. Using data envelopment analysis for evaluating flexible manufacturing systems in the presence of imprecise data. Int. J. Adv. Manuf. Technol. 35, 867-874.

Lebacq, T., Baret, P., Stilmant, D., 2015. Role of input self-sufficiency in the economic and environmental sustainability of specialised dairy farms. Animal 9, 544-552.

Mugera, A.W., 2013. Measuring technical efficiency of dairy farms with imprecise data: a fuzzy data envelopment analysis approach. Aust. J. Agric. Resour. Econ. 57, 501-520.

Müller, K., Holmes, A., Deurer, M., Clothier, B.E., 2015. Eco-efficiency as a sustainability measure for kiwifruit production in New Zealand. Journal of Cleaner Production 106, 333-342. 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.

Meul, M., Nevens, F., Verbruggen, I., Reheul, D., Hofman, G., 2007. Operationalising eco-efficiency in agriculture: the example of specialised dairy farms in Flanders. Progr. Ind. Ecol 4, 41-53.

Mu, W., Groen, E.A., van Middelaar, C.E., Bokkers, E.A.M., Hennart, S., Stilmant, D., de Boer, I.J.M., 2017. Benchmarking nutrient use efficiency of dairy farms: The effect of epistemic uncertainty. Agric. Sys.156, 25-33.

Mu, W., van Middelaar, C., Bloemhof, J., Oenema, J., de Boer, I., 2016. Nutrient balance at chain level: a valuable approach to benchmark nutrient losses of milk production systems. J. Clean. Prod.112, 2419-2428.

Picazo-Tadeo, A.J., Gómez-Limón, J.A., Reig-Martínez, E., 2011. Assessing farming eco-efficiency: a data envelopment analysis approach. J. Environ. Manage. 92, 1154-1164.

Ramilan, T., Scrimgeour, F., Marsh, D., 2011. Analysis of environmental and economic efficiency using a farm population micro-simulation model. Math. Comput. Simulation. 81, 1344-1352.

Seppäläa, J., Melanen, M., Mäenpää, I., Koskela, S., Tenhunen, J., Hiltunen, M.R., 2005. How Can the Eco-efficiency of a Region be Measured and Monitored? J. Ind. Ecol. 9, 117-130.

Stokes, J., Tozer, P., Hyde, J., 2007. Identifying efficient dairy producers using data envelopment analysis. J. Dairy Sci. 90, 2555-2562.

Shortall, O., Barnes, A., 2013. Greenhouse gas emissions and the technical efficiency of dairy farmers. Ecol. Indic. 29, 478-488.

Sefeedpari, P., Rafiee, S., Akram, A., 2012. Selecting energy efficient poultry egg producers: a fuzzy data envelopment analysis approach. Int. J. Appl. Oper. Res. 2, 77-88.

Srinivasa Raju, K., Nagesh Kumar, D., 2013. Fuzzy data envelopment analysis for performance evaluation of an irrigation system. Irrig. Drain. 62, 170-180.

Spillman, W., 1923. Application of the law of diminishing returns to some fertilizer and feed data. J. Far. Econ. 5, 36-52.

Thanawong, K., Perret, S., Basset-Mens, C., 2014. Eco-efficiency of paddy rice production in Northeastern Thailand: a comparison of rain-fed and irrigated cropping systems. J. Clean. Prod. 73, 204-217.

Urdiales, M.P., Lansink, A.O., Wall, A., 2016. Eco-efficiency among dairy farmers: the importance of socio-economic characteristics and farmer attitudes. Environ. Resour. Econ. 64, 559-574.

Van Middelaar, C., Berentsen, P., Dolman, M., De Boer, I., 2011. Eco-efficiency in the production chain of Dutch semi-hard cheese. Livest. Sci. 139, 91-99.

Wanke, P., Barros, C.P., Emrouznejad, A., 2016. Assessing productive efficiency of banks using integrated Fuzzy-DEA and bootstrapping: A case of Mozambican banks. Eur. J. Oper. Res. 249, 378-389.

Zhang, B., Bi, J., Fan, Z., Yuan, Z., Ge, J., 2008. Eco-efficiency analysis of industrial system in China: a data envelopment analysis approach. Ecol. Econ. 68, 306-316.

Chapter 6

General discussion

6.1 Introduction

Benchmarking dairy farming systems offers the option to identify best farm practices and to provide guidance for improving farm performance (Stokes et al., 2007; Dolman et al., 2014). Currently, benchmarking is hampered by the lack of an effective method that results in a set of indicators that is easily quantifiable and detects variations in environmental performance between farms. The aim of this thesis, therefore, was to develop a sound method to benchmark the environmental performance of dairy farming systems.

Benchmarking consists of three steps (see Figure 6.1). In the first step, the goal and scope of the (benchmarking) assessment are defined, and the boundary of the system is determined. The overall goal of benchmarking the environmental performance of dairy farming systems is to reduce their impact on the environment. The environmental impacts included throughout this thesis (scope) are climate change, acidification, eutrophication, energy use and land use. The system boundary is defined from cradle-to-farm-gate, implying that we included the main processes up to the moment milk leaves the farm-gate. The second step involves the selection of indicators to benchmark the dairy farming systems. In this step, we explored the value of environmental indicators derived from different methods (Chapter 2 and 3) and selected a set of indicators that can explain variations in environmental performance between dairy farming systems (Chapter 3). The third step is to assess and compare this performance. In this step, we investigated the importance of benchmarking farms with peers (Chapter 2, 4 and 5) and the impact of epistemic uncertainty of data on benchmarking results (Chapter 4 and 5).

In this chapter we discuss the overall findings of this thesis. In addition, research implications are elaborated, and future research opportunities are pointed out. This chapter is organised as follows. Section 2 discusses the value of different environmental indicators based on the results of this thesis. Section 3 discusses the issues linked to the selection of a set of indicators that can explain variations in environmental performance between dairy farming systems. Section 4 discusses the reasons why it is important to benchmark with peers. Section 5

discusses findings with regard to the impact of epistemic uncertainty of data on benchmarking results. In section 6, potential future research areas are pointed out, whereas in section 7 main conclusions are presented.
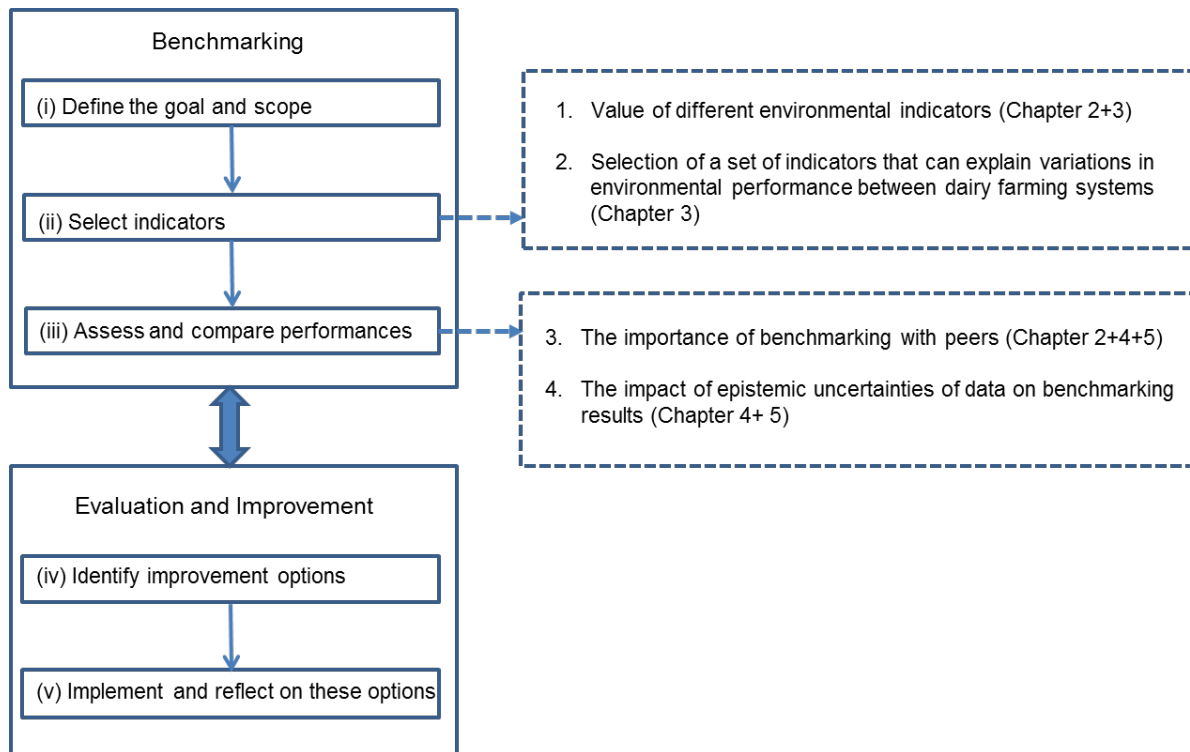

Figure 6.1. The conceptual framework for improving the environmental performance of dairy farming systems, adjusted from van der Vorst et al. (2013) and de Olde et al. (2017), and issues addressed in this thesis.

6.2 Value of different environmental indicators

An indicator suitable for benchmarking should be relevant, measurable, valid, timely and understandable (Bell and Morse,1999; Lockie et al., 2002). In the past decades, various indicators have been developed to assess the environmental performance of dairy farming systems. These indicators are derived from different methods, such as a nutrient balance (NB) approach or a life cycle assessment (LCA), and can examine different environmental impacts (e.g. climate change, biodiversity) at different levels (e.g. farm, chain) (Gerber et al., 2013).

When the benchmarking purpose is to raise the awareness of decision-makers on nutrient (mis)use and nutrient use (in)efficiency, indicators from an NB approach can be adopted,

because they are easy to quantify (i.e. measurable) and easy to explain to decision-makers (i.e. understandable) (Oenema et al., 2003; Halberg et al., 2005; Gourley et al., 2012). For example, the Dutch policy instrument Annual Nutrient Cycling Assessment (ANCA, in Dutch: Kringloopwijzer) has been developed and introduced to monitor and stimulate improvement in nitrogen (N) and phosphorus (P) use efficiency on dairy farms (Daatselaar et al., 2015). Another example is the EU Nitrogen Expert Panel (2015), who suggests to interpret N use efficiency in relation to productivity (N output) and N surplus ((N input − N output)/ha) to improve N use efficiency in the food chain.

Indicators from an NB approach, however, are often quantified at farm level and do not take into account the nutrient losses occurring during off-farm processes, such as the production of purchased feed, which can lead to biased benchmarking results. Chapter 2, for example, shows that the ranking of farms based on NB indicators at farm level can differ from the ranking based on NB indicators at chain level. This means that farms that are relatively efficient at farm level can be less efficient at chain level, and vice versa. When off-farm nutrient losses are relatively more important compared to on-farm losses (i.e. farms heavily rely on purchased feed), and when changes in off-farm losses per unit change in on-farm losses differ between farms (i.e. farms differ in management), chain-based NB indicators are recommended. As a first step, decision-makers are advised to start with a farm-based NB. For regions where dairy farms are relatively self-sufficient in terms of feed, such as in New Zealand and Ireland, a farm-based NB can be used for benchmarking their nutrient losses. Also for farms with a similar amount of purchased feed per cow per year and a similar feed composition, a farm-based NB can be used for benchmarking. Only if farms differ largely in the amount and/or type of purchased feed, the farm-based NB should be extended to a chain-based NB. Nevertheless, when exploring improvement options, the whole chain should be taken into account. In addition, decision-makers should be aware of the disadvantage of indicators from an NB approach. An NB approach does not specify the type of losses and their associated

environmental impacts (e.g. acidification potential) and neglects certain environmental impact categories, such as energy use or biodiversity loss (Thomassen and De Boer, 2005).

Contrary to an NB approach, an LCA does specify the types of losses, as well as the potential associated environmental impacts related to those losses. An LCA, moreover, can address more environmental impacts than an NB, such as fossil energy depletion, water deprivation or biodiversity loss. Studies, however, suggest that collection of data required to perform an LCA appears difficult and time consuming (Thomassen and De Boer, 2005). The global warming potential per kg of milk, for example, also known as carbon footprint, is an LCA indicator that assesses the potential impact on climate change related to the production of one kg of milk, by summing emissions of greenhouse gases (i.e. CO_2, CH_4 and N_2O) along the production chain. Over the past few years, the number of LCA studies on milk production have increased enormously (e.g. Cederberg, 1998; Thomassen and De Boer, 2005; Yan et al., 2011). Environmental impacts, such as climate change, acidification, eutrophication, land use and energy use, are often examined in these LCA studies. Environmental problems, such as human and eco-toxicity, water deprivation and biodiversity loss, are less examined for two reasons. First, human and eco-toxicity, for example, are often not quantified because accurate data about pesticides or heavy metal use are lacking (Yan et al., 2011). Similarly, knowledge about water use (i.e. rain water and irrigation water) during crop cultivation across the world is often unknown, which hampers computation of a water footprint. Assessing the impact of livestock production on biodiversity loss is complex, as diversity needs to be considered at the level of genetics, species and ecosystem. There is currently no method that can provide a common unit to address the aforementioned three levels, across the production chain and the complexity within each level (Atkinson et al., 2000; Yan et al., 2011).

In most LCA studies of milk, the system boundary is defined from cradle-to-farm gate (de Boer, 2003). This boundary is generally justified because the agricultural stages of the milk production chain, including the production of feed and keeping of dairy cows and young stock explain the majority of the environmental impacts (De Vries and de Boer, 2010; Djekic et al.,

2014). In this thesis, we also defined the system boundary from cradle-to-farm-gate, so post-farm stages, such as milk processing, product storage, retail or final consumption, were excluded. The contribution of post-farm-gate stages to processing of milk into, for example, yoghurt or cheese is relatively small for most impacts, except for energy use, where the contribution of post-farm stages can be quite substantial. Van Middelaar et al. (2011), for example, found that the post-farm stages cheese-making, storage, packaging and retail in total contributed about 10% to GHG emissions, 2% to land use, but 31% to energy use. Benchmarking of diversity of processes used in post-farm stages of milk, therefore, is especially important in terms of energy use.

Despite the relatively low contribution of post-farm stages to, for example, climate change or land use, it should be kept in mind that improving the efficiency of conversion of milk into milk products, such as cheese or yoghurt, could significantly reduce the environmental impact of those products (Flysjö 2012; Djekic et al., 2014).

We have chosen to benchmark dairy farming systems based on product-based indicators, i.e. indicators that determine the environmental impact per kg product. Using product-based indicators is justifiable for so-called global environmental problems, such as climate change, fossil energy depletion or fossil phosphorus depletion (De Boer, 2003). Emissions of acidifying gases or eutrophying substrates, however, also have a strong impact on the local environment. For these environmental impacts, decision-makers should not only rely on product-based indicators, but also explore the regional impact of farming systems. They should, for example, assess the impact of all farms and their distribution in a region to examine if there are regional differences in farm density and the tolerance of the local ecosystem to the environmental emissions. To prevent a too high environmental impact in a specific region, the government could introduce regulations, such as total production volume restrictions or maximum number of animals per ha.

6.3 Selection of a set of indicators that can explain variations in environmental performance of farming systems

In recent years, the carbon footprint has been used as 'the indicator' for assessing the environmental impact of dairy products by stakeholders, such as FrieslandCampina in the Netherlands and Arla Foods in Denmark and Sweden (Oenema, 2013). In addition, a large number of scientific studies have focussed on reducing greenhouse gas (GHG) emissions per unit of milk only (Casey and Holden, 2005; Flysjö, 2011; Thoma et al., 2013; Finnegan et al., 2015). One the one hand, climate change is considered to be a very important environmental challenge to the planet (Steffen et al., 2015). On the other hand, policy-makers and industries relying on the carbon footprint as the sole environmental indicator run a risk of overlooking other environmental issues, such as human and eco-toxicity, acidification, and water deprivation (Laurent et al., 2012).

A single indicator cannot reflect the complexity of environmental impacts related to a product or product system (Berger and Finkbeiner, 2011). A set of indicators, therefore, is needed to have a correct measurement of the environmental performance of that system (Lyytimäki and Rosenström, 2008). Exploring correlations between various indicators from different assessment methods can help to reduce unnecessary indicators, and to identify this set of indicators (Lebacq et al., 2013). Previous studies have mainly focused on correlations between LCA indicators (Berger and Finkbeiner, 2011; Laurent et al., 2012; Röös et al., 2013), while no study examined correlations between environmental indicators from different methods within dairy production. We, therefore, explored correlations between NB and LCA indicators to identify an effective set of indicators that can be used as a proxy for benchmarking the environmental performance of dairy farming systems (see Chapter 3). All the indicators are computed with the system boundary from cradle-to-farm-gate. Results show that a set of indicators, consisting of N surplus, P surplus, land use and energy use, can be used as a proxy to benchmark the environmental performance of dairy farming systems, and to represent global warming potential (GWP), acidification potential, freshwater eutrophication potential,

and marine eutrophication potential. The selection of this set of indicators, however, could have been influenced by the calculation methods that were used. Data availability influenced the methods chosen to quantify some of the emissions, and, therefore, also influenced the determined variations in environmental impacts between farms. Calculations on enteric CH_4 emissions, for example, were based on guidelines of the Intergovernmental Panel on Climate Change, or IPCC (IPCC, 2006). These guidelines are often used to assess GHG emissions of the dairy sector. The IPCC classifies the methodological approach into three tiers based on the quantity of information needed and the degree of analytical complexity (IPCC, 2006). Although Tier 3 methods are most accurate, they generally require extensive farm specific information. Tier 2 methods are less accurate, but they are most commonly adopted when comparing results across regions and farming systems, because of their relatively lower data requirement. In this study, enteric CH_4 emissions were estimated based on IPCC (2006) Tier 2. We found that the variation in CH_4 emission per liter of milk between farms was small, which related to the fact that emissions were assumed equal to 6.5% of the gross energy intake of the dairy herd. Gross energy intake of the herd was estimated based on farm-specific data about herd composition and milk production, and generic data on energy requirements for growth, maintenance, activity, and milk production, and on digestibility of the feed. The effect of differences in dietary composition of the herd on CH_4 emission, therefore, was not included in the calculation method. As a result, differences in enteric CH_4 emissions among farms reflected differences in efficiency of milk production only, and GWP did not end in the final selected set of indicators, because of high correlations between GWP and N surplus and energy use. If information on, e.g., dietary composition would be available, CH_4 emission per liter of milk could have been assessed with a more detailed approach (e.g. Tier 3), and GWP might have ended up in the final set of indicators. When information on the exact feed composition is not available, as is common in practice, however, differences between CH_4 emission mainly relate to differences in efficiency of milk production, which are captured also by other indicators, such as energy use.

N surplus, P surplus, land use and energy use are a good set of indicators to start benchmarking dairy farming systems across the world. This set of indicators can be monitored and collected in a time and cost-effective way (i.e. measurable). In addition, these indicators can be interpreted easily by decision-makers (i.e. understandable), and their application can contribute to the consistency and comparability of results across regions and farming systems. However, the selected set of indicators cannot provide information on other important environmental impacts, such as biodiversity and water use. The work in Chapter 3, therefore, can be seen as a first attempt to derive a set of indicator to benchmark dairy farming systems. In the future, the same principle can be applied to derive a set of indicators that can provide a more comprehensive assessment of the environmental performance of dairy farming systems.

6.4 The importance of benchmarking with peers

When benchmarking the environmental performance of farms, we should acknowledge that differences in performances between farms might result from unmanageable factors, such as differences in, for example, soil type or climate. It is, therefore, important to compare farms with so-called peers, i.e. farms that are comparable in terms of characteristics, such as size, intensity and site-specific circumstances (de Snoo, 2006; Dolman et al., 2014). Dairy farms in Europe are heterogeneous in management. Most farms in the Netherlands, for example, have a high stocking density and in addition to on-farm produced feed, substantive amounts of concentrates are purchased. Most farms in Ireland, however, are grass-based extensive farms. Their main input consists of purchased mineral fertilizer. Benchmarking Dutch and Irish farms based on their environmental performance can lead to biased conclusions, because of inherent differences between systems. In Chapter 4, we therefore clustered farms based on 5 characteristics (i.e. grazing hours, soil type, concentrates per cow per year, milk production per cow per year and milk production per ha) before benchmarking the nutrient use efficiency of these farms. By doing so, farms are compared with their peers, allowing identification of specific improvement options based on incremental changes in, for example, feed composition or manure management. A discussion about the environmental performances of contrasting

farms, such as large-scale conventional farms, organic farms and pure-graze dairy farming systems, could contribute to the development of so-called system innovations, such as a complete new farm design. The "Rondeel" in the Netherlands, for example, is a new housing design for laying hens that was developed by egg farmers together with poultry scientist, retailers, animal welfare leaders and consumer representatives (van Niekerk and Reuvekamp, 2011). Evaluation and comparison of existing systems has been an important step in the development of the Rondeel (van Niekerk and Reuvekamp, 2011).

6.5 The impact of epistemic uncertainties of data on benchmarking results

Milk production involves complex production processes, so epistemic uncertainty can arise from measurement errors, such as errors in determining the N fixation by clover, errors in feed intake estimates of cows or errors in estimates of manure composition (Oenema et al., 2015). Data used to calculate the environment performance of dairy farming systems, therefore, can contain high levels of epistemic uncertainty (Sefeedpari et al., 2012; Mugera, 2013). Most studies that assess the environmental performance of dairy farming systems do not cope with data uncertainty (Schröder, 2003; Nevens, 2006; Powell, 2010). A few studies did incorporate data uncertainty, but they did not examine the impact of epistemic uncertainty on benchmarking results (Mulier et al., 2003; Basset-Mens et al., 2009; Chen et al., 2014; Oenema et al., 2015). In Chapter 4, we benchmarked the N use efficiency of farms and found that ranking of farms is not possible when the epistemic uncertainty of parameters is large and differences in N use efficiency are small. When benchmarking farms, it is therefore important to account for epistemic uncertainty in order to avoid biased benchmarking results.

We furthermore identified which parameters explained most of the variance in N use efficiency, and subsequently reduced the uncertainty of these parameters to examine if this would improve our benchmarking results. We found that the most influential parameters differed per system. For the concentrate-based system, input of concentrates and roughage and output of roughage were most important; whereas for the grass-based system, the input of mineral

fertiliser was most important. Results showed that when epistemic uncertainty of the most influential input parameters are reduced to 1%, a rather optimistic scenario for some of the parameters, benchmarking results can be improved significantly. When benchmarking farming systems and implementing policy instruments (e.g. ANCA), decision-makers should be aware of those influential input parameters. They can draw more reliable conclusions on the performance of the farms by measuring those parameters more precisely.

To benchmark the environmental performance of dairy farming systems, we need to account for data uncertainty of multiple indicators. In Chapter 5, therefore, we demonstrated how to use fuzzy data envelopment analysis (DEA) to incorporate uncertainties of multiple indicators to benchmark dairy farming systems. In this chapter, we also included an economic indicator, because it is important for farms to stay economically viable while improving their environmental performance. Based on the ratio between the economic value and the environmental impact added by a system, its eco-efficiency can be defined (Keffer and Shimp 1999; Seppälä et al., 2005). In this case, the eco-efficiency was based on N surplus, P surplus, land use, energy use and gross margin. Compared to standard DEA, which is deterministic and does not enable to account for uncertainty, fuzzy DEA does take uncertainty into account by expressing all parameters as ranges instead of single crisp numbers (Emrouznejad et al., 2014; Egilmez et al., 2016). We assumed an uncertainty range of -20% to +20% of the observed level per indicator to determine the lower and upper bound of each indicator. By increasing the uncertainty of the parameters stepwise from the lowest (no uncertainty) up to the highest (20% uncertainty) level, fuzzy DEA provides benchmarking results for different levels of uncertainty. This enables decision-makers, for example, to choose the proper uncertainty level to draw conclusions. Based on the highest level of uncertainty, ranking of farms based on their eco-efficiency differed from ranking of farms based on the lowest level of uncertainty. In addition, the set of peers differed (i.e., the farms that could serve as an example for others), and the level of improvement potential per parameter per farm was represented by a range, rather than by a single value. Aforementioned findings show that decision-makers

should be aware that standard DEA can result in misleading and inaccurate conclusions when data used to calculate the indicators contain uncertainty. By taking uncertainty into account during the quantification processes, fuzzy DEA can contribute to increasing the reliability of results and prevent biased conclusions. Based on the findings of Chapter 3 and 4, we can conclude that benchmarking farms based on indicators calculated from parameters that contain epistemic uncertainty can be misleading and inaccurate. When benchmarking farms, especially as a policy instruments (e.g. ANCA), therefore, decision-makers should acknowledge the effect of epistemic uncertainty. Instead of using a single value as the reference value for giving a penalty, for example, this reference value should be based on a range of values that accounts for the epistemic uncertainty of input parameters.

6.6 Future research

In this research, we mainly focused on product-based indicators. However, we did not incorporate competition for natural resources, such as land, between humans and animals. By grazing on marginal land, i.e. land that is less suitable or even unsuitable for crop production because of rainfall, temperature or poor terrain limitations, dairy cows can convert inaccessible nutrients to human edible products (Eisler et al., 2014; van Zanten et al., 2016). Overlooking competition for land between humans and animals might lead to biased conclusions. For example, in our study we found concentrate-based systems being more efficient than grass-based systems. In case of cows grazing on marginal land, however, the grass-based systems are preferred over concentrate-based systems in terms of land use efficiency, when accounting for competition for land between humans and animals. Future studies, therefore, are suggested to consider the suitability of land for growing crops that can be consumed directly by humans in order to account for the competition for land between humans and animals (Van Zanten et al., 2016).

In addition, this thesis focused on European specialised dairy farming systems, where an important function of dairy cattle is to produce marketable products, like milk and meat. In many

other dairy farming systems, however, cattle fulfil multiple functions. In smallholder dairy systems, cattle often are kept not only to produce milk and meat, but also to produce fertilizer and draught power to sustain crop production. Cattle, furthermore, can have an insurance and financing function in these systems. Grazing ruminants, moreover, can contribute to non-provisioning ecosystem services, such as preservation and enhancement of biodiversity or preservation of the landscape (Ripoll-Bosch et al., 2013; Weiler et al., 2014). Due to this multi-functionality, not all systems can be analysed in the same way as the specialised systems included in this study (Weiler et al., 2014). Neglecting the multi-functionality of cattle in, e.g. smallholder systems, will result in an overestimation of the impacts per unit of product produced (Weiler et al., 2014). Future researches, therefore, are suggested to take into account of multi-functionality in the benchmarking processes for non-specialised dairy systems.

Finally, this thesis mainly focused on environmental performance of dairy farming systems. However, the importance of the social and economic performance of those systems should not be neglected. There are potential trade-offs between environmental indicators, and social or economic indicators, such as trade-offs between the carbon footprint of a product and animal welfare indicators, or between biodiversity impacts and net farm income (Steffan-Dewenter et al., 2007). Future studies are suggested to explore the correlation between environmental, social or economic indicators.

6.7 Conclusions

This thesis contributes to the development of a sound method to benchmark the environmental performance of dairy farming systems by investigating ways to identify an effective set of indicators and to handle the effect of data uncertainty on benchmarking results.

- To benchmark dairy farming systems, a nutrient balance (NB) at farm level can be used if differences in on-farm losses between systems are large, and off-farm losses are relatively unimportant. To benchmark individual farms, an NB at farm level can be used

if changes in off-farm losses per unit change in on-farm losses are similar across farms. A chain level nutrient balance of a sample set, however, is required to verify these conditions.

- Data availability is a major challenge for benchmarking. Exploring correlations between environmental indicators from different environmental impact assessment methods can facilitate decision-makers to derive an effective set of indicators that can be used as proxies for benchmarking purposes.

- Nitrogen surplus, phosphorous surplus, energy use and land use along the chain can be used to represent a broader set of indicators, including also global warming potential, acidification potential, freshwater eutrophication potential and marine eutrophication potential. This 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. More detailed methods to assess the global warming potential of systems could increase the importance of this indicator. Other important environmental impacts, such as biodiversity and water use, however, should not be overlooked.

- Epistemic uncertainty of data leads to unreliable benchmarking results and sometimes makes benchmarking impossible. Identifying and reducing the epistemic uncertainty of major parameters can significantly improve benchmarking results. The parameters that are most important in terms of epistemic uncertainty (i.e., explain most of the output variance), however, can vary among dairy farming systems. For the concentrate-based systems in the Netherlands, for example, input of concentrates and roughage and output of roughage were the most important parameters in terms of epistemic uncertainty.

- Decision-makers should acknowledge the effect of epistemic uncertainty to avoid biased conclusion. When setting up reference values for penalty, for example, this value should be based on a range rather than a single value in order to account for epistemic uncertainty.

- Using product-based indicators is justifiable for so-called global environmental impacts, such as climate change, fossil energy or fossil phosphorus depletion. In case of environmental impacts with a local importance, such as eutrophication and acidification, however, decision-makers should not only rely on product-based indicators. To prevent a too high environmental impact in a specific region, decision-makers could introduce regulations, such as total production volume restrictions or maximum number of animals per ha.

- Future researches on benchmarking farming systems are suggested to consider the suitability of land for growing crops that can be consumed directly by humans in order to account for the competition for land between humans and animals. Moreover, future researches should account for multi-functionality when benchmarking non-specialised dairy systems, and explore correlation between environmental, social or economic indicators.

6.8 References

Atkinson, S.F., Bhatia, S., Schoolmaster, F.A., Waller, W.T., 2000. Treatment of biodiversity impacts in a sample of US environmental impact statements, Impact Assess Proj Apprais, 18, 271-282,

Basset-Mens, C., Kelliher, F.M., Ledgard, S., Cox, N., 2009. Uncertainty of global warming potential for milk production on a New Zealand farm and implications for decision making. Int. J. Life. Cycle. Assess.14, 630-638.

Berger M., Finkbeiner, M., 2011. Correlation analysis of life cycle impact assessment, indicators measuring resource use. Int. J. Life Cycle Assess. 16, 74–81.

Bell S, Morse S., 1999. Sustainability Indicators: Measuring the Immeasurable. Earthscan: London.

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, Goteborg, Sweden.

Casey, J.W., Holden, N. M., 2005. Analysis of greenhouse gas emissions from the average Irish milk production system. Agric. Syst. 86, 97-114.

Chen, X., Corson, M.S., 2014. Influence of emission-factor uncertainty and farm-characteristic variability in LCA estimates of environmental impacts of French dairy farms. J. Clean. Prod. 81, 150-157.

De Boer, I.J.M., 2003. Environmental impact assessment of conventional and organic milk production. Livest Prod Sci. 80, 69–77.

Daatselaar, C.H., Reijs, J.R., Oenema, J., Doomewaard, G.J., Aarts, H.F., 2015. Variation in nutrient use efficiencies on Dutch dairy farms. J. Sci. Food Agric. 95,15.
De Olde, E.M., Bokkers, E.A.M., De Boer, I.J.M., 2017. The Choice of the Sustainability Assessment Tool Matters: Differences in Thematic Scope and Assessment Results. Ecol. Econ. 136, 77-85.

Djekic, I., Miocinovic, J., Tomasevic, I., Smigic, N., Tomic, N., 2014. Environmental life-cycle assessment of various dairy products. J. Clean. Prod., 68, 64-72

Dolman, M., Sonneveld, M., Mollenhorst, H., De Boer, I., 2014. Benchmarking the economic, environmental and societal performance of Dutch dairy farms aiming at internal recycling of nutrients. J. Clean. Prod. 73, 245-252.

De Snoo, G.R., 2006. Benchmarking the environmental performances of farms. Int. J. LCA 11, 22-25.

De Vries, M., De Boer, I.J., 2010. Comparing environmental impacts for livestock products: A review of life cycle assessments. Livest. Sci. 128, 1-11.

Emrouznejad, A., Tavana, M., Hatami-Marbini, A., 2014. Performance measurement with fuzzy data envelopment analysis. Studies in Fuzziness and Soft Computing, Springer-Verlag, Berlin Heidelberg.

Egilmez, G., Gumus, S., Kucukvar, M., Tatari, O., 2016. A fuzzy data envelopment analysis framework for dealing with uncertainty impacts of input–output life cycle assessment models on eco-efficiency assessment. J. Clean. Prod. 129, 622-636.

Eisler, M.C., Lee, M.R.F., Tarlton, J.F., Martin, G.B., Beddington, J., Dungait, J.A.J., Greathead, H., Liu, J., Mathew, S., Miller, H., Misselbrook, T., Murray, P., Vinod, V.K., Van Saun, R., Winter, M., 2014, Steps to sustainable livestock. Nature, 507, 32–34.

EU Nitrogen Expert Panel, 2015. Nitrogen Use Efficiency (NUE) – An Indicator for the Utilization of Nitrogen in Agriculture and Food Systems. Wageningen University, Netherlands.

Flysjö, A.M.,2012. Greenhouse Gas Emissions in Milk and Dairy Product Chains: Improving the Carbon Footprint of Dairy Products, PhD thesis, Institut for Agroøkologi. Aarhus Universitet, Aarhus Universitet, Aarhus, Denmark.

Finnegan, W., Goggins, J., Clifford, E., Zhan, X., 2015. Global warming potential associated with dairy products in the Republic of Ireland. J. Clean. Prod. In press.

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.I., 2012. Farm-scale nitrogen, phosphorus, potassium and sulphur balances and use efficiencies on Australian dairy farms. Anim. Prod. Sci. 52, 929–944.

Gerber, P.J., Steinfeld, H., Henderson, B., Mottet, A., Opio, C., Dijkman, J., Falcucci, 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.

Halberg, N., van der Werf, H., Basset-Mens, C., Dalgaard, R., de Boer, I.J.M., 2005. Environmental assessment tools for the evaluation and improvement of European livestock production systems. Livest. Prod. Sci. 96, 33–50.

IPCC, 2006. Intergovernmental Panel on Climate Change. Guidelines for national greenhouse gas inventories. Volume 4: agriculture, forestry and other land use. Prepared by the National Greenhouse Gas Inventories Program. Eggleston HS, Buendia L, Miwa K, Ngara T, Tanabe K (eds). IGES, Japan

Keffer, C., Shimp, R., Lehni, M., 1999. Eco-efficiency indicators and reporting. World Business Council for Sustainable Development (WBCSD), London.

Laurent, A., Olsen, S.I., Hauschild, M.Z., 2012. Limitations of carbon footprint as indicator of environmental sustainability. Environ. Sci. Technol. 46,4100–4108.

Lyytimäki, J., Rosenström, U., 2008. Skeletons out of the closet: effectiveness of conceptual frameworks for communicating sustainable development indicators. Sustain. Dev.16, 301-313.

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.

Lebacq, T., Baret, P.V., Stilmant, D., 2013. Sustainability indicators for livestock farming. A review. Agron. Sustain. Dev., 33, 311-327.

Mugera, A.W., 2013. Measuring technical efficiency of dairy farms with imprecise data: a fuzzy data envelopment analysis approach. Aust. J. Agric. Resour. Econ. 57, 501-520.

Mulier, A. , Hofman, G. , Baecke, E. , Carlier, L. , De Brabander, D. , De Groote, G. , 2003. A methodology for the calculation of farm level nitrogen and phosphorus balances in Flemish agriculture. Eur. J. Agron, 20, 45–51.

Nevens, F., Verbruggen, I., Reheul, D., Hofman. G., 2006. Farm gate nitrogen surpluses and nitrogen use efficiency of specialized dairy farms in Flanders: Evolution and future goals. Agric. Syst. 88, 142–155.

Oenema, J., 2013. Transitions in nutrient management on commercial pilot farms in the Netherlands. PhD thesis, Wageningen University, Wageningen, The Netherlands.

Oenema, J., Burgers, S., van Keulen, H., van Ittersum, M., 2015. Stochastic uncertainty and sensitivities of nitrogen flows on dairy farms in The Netherlands. Agric. Syst. 137, 126-138.

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.

Ripoll-Bosch, A., De Boer, I.J.M., Bernues, A., Vellinga, T.V., 2013. Accounting for multifunctionality of sheep farming in the carbon footprint of lamb: a comparison of three contrasting Mediterraneansystems. Agric. Syst.116, 60-68.

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.

Steffan-Dewenter, I., Kessler, M., Barkmannc, J., Bosa, M., Buchorig, D., Erasmih, S., Fausth, H., Geroldh, G., Glenke, K., Gradsteind, S.R., Guhardjai, E., Harteveldd, M., Herteld, D., Höhna, P., Kappash, M., Köhlerh, S., Leuschnerd, C., Maertensj, M., Marggrafe, R., Migge-Kleiank, S., Mogeai, J., Pitopangl, R., Schaeferk, M., Schwarzem, S., Spornd, S.G., Steingrebek, A., Tjitrosoedirdjoi, S.S., Tjitrosoemitoi, S., Tweleh, A., Weberh, R., Woltmannk, L., Zellerm, M.N., Tscharntke, T., 2007. Tradeoffs between income, biodiversity, and ecosystem functioning during tropical rainforest conversion and agroforestry intensification. PNAS, 104, 4973–4978.

Schröder, J.J., Aarts, H.F.M., Ten Berge, H.F.M., Van Keulen, H., Neeteson, J.J., 2003. An evaluation of whole-farm nitrogen balances and related indices for efficient nitrogen use. Eur. J. Agron. 20,33–44.

Seppälä, J., Melanen, M., Mäenpää, I., Koskela, S., Tenhunen, J., Hiltunen, M.R., 2005. How Can the Eco-efficiency of a Region be Measured and Monitored? J. Ind. Ecol. 9, 117-130.

Stokes, J., Tozer, P., Hyde, J., 2007. Identifying efficient dairy producers using data envelopment analysis. J. Dairy Sci. 90, 2555-2562.

Sefeedpari, P., Rafiee, S., Akram, A., 2012. Selecting energy efficient poultry egg producers: a fuzzy data envelopment analysis approach. Int. J. Appl. Oper. Res. 2, 77-88.

Steffen, W., Richardson, K., Rockström, J., Cornell, S.E., Fetzer, I., Bennett, E.M., Biggs, R., Carpenter, S.R., de Vries, W., de Wit, C.A., Folke, C., Gerten, D., Heinke, J., Mace, G.M., Persson, L.M., Ramanathan, V., Reyers, B., Sörlin, S., 2015. Planetary boundaries: guiding human development on a changing planet. Science, 347, 6223.

Thoma, G., Popp, J., Nutter, D., Shonnard, D., Ulrich, R., Matlock, M., Kim, D.S., Neiderman, Z., Kemper, N., East, C., Adom, F., 2013. Greenhouse gas emissions from milk production

and consumption in the United States of America: A cradle-to-grave life cycle assessment circa 2008. Int. Dairy J., 31, S3-S14

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.

Van der Vorst, J.G., Peeters, L., Bloemhof, J.M., 2013. Sustainability Assessment Framework for Food Supply Chain Logistics: Empirical Findings from Dutch Food Industry. Int. J. Food. Sys. Dynam., 130-139.

Van Middelaar, C.E., Berentsen, P.B.M., Dolman, M.A., de Boer, I.J.M. 2011. Eco-efficiency in the production chain of Dutch semi-hard cheese. Livest. Sci., 139, 91-99.

Van Niekerk, T., Reuvekamp, B., 2011. Rondeel, a new housing design for laying hens. Lohmann Information, 46, 25.

van Zanten, H.H.E., Mollenhorst, H., Klootwijk, C.W., van Middelaar, C.E., de Boer, I.J.M. 2016. Global food security: land use efficiency of livestock systems. Int. J. Life Cycle Assess. 21, 747-758.

Weiler, V., Udo, H.M.J., Viets, T., Crane, T.A., De Boer, I.J.M., 2014. Handling multi-functionality of livestock in a life cycle assessment: the case of smallholder dairying in Kenya Curr. Opin. Environ. Sustain. 8, 29-38.

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.

Summary

Milk production causes pressure on the environment by its emissions to air, water, and soil, that contribute to, for example, climate change, eutrophication, acidification, and loss of biodiversity. Furthermore, milk production competes increasingly for scarce resources, such as land, water and fossil energy. As the population grows and becomes wealthier, milk production is expected to increase further. Aforementioned environmental consequences related to milk production, therefore, are likely to amplify.

Benchmarking dairy farming systems offers the opportunity to identify best farm practices and to provide guidance for improving farm performance. Currently, benchmarking is hampered by the lack of an effective method that results in a set of indicators that is easily quantifiable and detects variations in environmental performance between farms. The aim of this thesis, therefore, was to develop a sound method to benchmark the environmental performance of dairy farming systems.

The first challenge in benchmarking the environmental performance of dairy farming systems is to select a set of indicators that are relevant, measurable, valid, timely and understandable. Environmental indicators can be derived from various approaches, including a nutrient balance (NB) approach and a life cycle assessment (LCA). An NB is generally applied at farm level, and yields indicators that are relatively easy to quantify and communicate. In **Chapter 2**, we tested whether an NB at chain level (i.e. including on-farm processes and the production of purchased feed) is more suited to benchmark nutrient losses of milk production systems or individual farms than an NB at farm level. Results showed that to benchmark dairy farming systems, an NB at farm level can be used if differences in on-farm losses between systems are large, and off-farm losses are relatively unimportant. To benchmark individual farms, an NB at farm level can be used if changes in off-farm losses per unit change in on-farm losses are similar across farms. A chain level balance of a sample set, however, is required to verify these conditions. As a first step, decision-makers are advised

to start with a farm-based NB. For regions where dairy farms are relatively self-sufficient in terms of feed, such as in New Zealand and Ireland, a farm-based NB can be used for benchmarking nutrient losses. Also for farms with a similar amount of purchased feed per cow per year and a similar feed composition, a farm-based NB can be used. Only if farms differ largely in the amount and/or type of purchased feed, the farm-based NB should be extended to a chain-based NB.

Environmental indicators derived from an NB approach, however, generally do not specify the type of losses, nor the environmental impacts associated with those losses, such as the impact on climate change or acidification. In addition, an NB approach neglects certain environmental impact categories (e.g. energy use). Contrary to the indicators from an NB approach, indicators from an LCA do specify the types of losses, as well as the potential associated environmental impacts related to those losses. However, LCA requires extensive data information which can be difficult to collect. In **Chapter 3**, correlations between eight commonly used NB and LCA indicators were explored to identify an effective set of indicators that can be used as a proxy for benchmarking the environmental performance of dairy farming systems. Indicators included were nitrogen (N) surplus, phosphorous (P) surplus, land use, fossil energy use, global warming potential (GWP), acidification potential (AP), freshwater eutrophication potential (FEP) and marine eutrophication potential (MEP). All the indicators are computed with the system boundary from cradle-to-farm gate. Based on multiple regression analysis and canonical correlation analysis, we found that a set of indicators, consisting of N surplus, P surplus, land use and energy use, can be used as a proxy to benchmark the environmental performance of dairy farms, and represent also MEP, FEP, AP and GWP. N surplus, P surplus, land use and energy use are a good set of indicators to start benchmarking dairy farming systems across the world, because they can be monitored and collected in a time and cost-effective way, and can be interpreted easily by decision-makers. More detailed methods to assess the GWP of systems, however, could

increase the importance of this indicator. Other important environmental impacts, such as biodiversity and water use, however, should not be overlooked.

The second challenge in benchmarking the environmental performance of dairy farming systems is to cope with data uncertainties. Milk production involves complex production processes, and data used to calculate the environment performance of dairy farming systems can contain high levels of epistemic uncertainty. Epistemic uncertainty can arise, for example, from measurement errors in N fixation by clover, errors in feed intake estimates of cows, or errors in estimates of manure composition. In **Chapter 4,** we evaluated the effect of epistemic uncertainty on benchmarking the N use efficiency of dairy systems. Results showed that ranking of farms based on this single indicator is not possible when the epistemic uncertainty of parameters is large and differences in N use efficiency are small. We, furthermore, found that the most influential parameters differed per system. For concentrate-based systems, input of concentrates and roughage and output of roughage were most important, whereas for grass-based systems, the input of mineral fertiliser was most important. Reducing epistemic uncertainty of the most influential input parameters improved benchmarking results significantly. It was concluded that to avoid biased benchmarking results, it is important to account for epistemic uncertainty.

To benchmark the environmental performance of dairy farming systems, we need to account for data uncertainty of multiple indicators. In **Chapter 5**, therefore, we demonstrated how to use fuzzy data envelopment analysis (DEA) to incorporate uncertainties of multiple indicators to benchmark dairy farming systems. In this chapter, we also included an economic indicator, because it is important for farms to stay economically viable while improving their environmental performance. Based on the ratio between the economic value and the environmental impact added by a system, its eco-efficiency can be defined. In this case, the eco-efficiency was based on N surplus, P surplus, land use, energy use and gross margin. Compared to standard DEA, which is deterministic and does not enable to account for uncertainty, fuzzy DEA does take uncertainty into account by expressing all parameters as

ranges instead of single crisp numbers. We assumed an uncertainty range of -20% to +20% of the observed level per indicator to determine the lower and upper bound of each indicator. By increasing the uncertainty of the parameters stepwise from the lowest (no uncertainty) up to the highest (20% uncertainty) level, fuzzy DEA provides benchmarking results for different levels of uncertainty. This enables decision-makers, for example, to choose the proper uncertainty level to draw conclusions. Based on the highest level of uncertainty, ranking of farms based on their eco-efficiency differed from ranking of farms based on the lowest level of uncertainty. In addition, the set of peers differed (i.e., the farms that could serve as an example for others), and the level of improvement potential per parameter per farm was represented by a range, rather than by a single value. Aforementioned findings show that decision-makers should be aware that standard DEA can result in misleading and inaccurate conclusions when data used to calculate the indicators contain uncertainty. By taking uncertainty into account during the quantification processes, fuzzy DEA can contribute to increasing the reliability of results and prevent biased conclusions.

In **Chapter 6**, challenges for developing a sound method to benchmark the environmental performance of dairy farming systems are discussed. The discussion revealed that data availability is a major challenge for benchmarking. Exploring correlations between environmental indicators from different environmental impact assessment methods, therefore, can facilitate decision-makers to derive an effective set of indicators that can be used as proxies for benchmarking. In addition, decision-makers should acknowledge the effect of epistemic uncertainty on benchmarking results. They should identify and reduce the epistemic uncertainty of major parameters to improve the accuracy of the benchmarking results. Moreover, when setting up reference values for penalty, for example, this value should be based on a range rather than a single value in order to account for epistemic uncertainty.

Chapter 6 also provides some future research opportunities. This thesis focused mainly on product-based indicators. However, caution should be made in case of environmental

impacts with a local importance, such as eutrophication and acidification. To prevent a too high environmental impact in a specific region, decision-makers could introduce regulations, such as total production volume restrictions or maximum number of animals per ha. Furthermore, we did not incorporate competition for natural resources, such as land, between humans and animals. Overlooking food-feed competition might lead to biased conclusions. For example, in our study we found concentrate-based systems being more efficient than grass-based systems. In case of cows grazing on marginal land, however, the grass-based systems are preferred over concentrate-based systems in terms of land use efficiency, when accounting for food-feed competition. Future studies, therefore, are suggested to consider the suitability of land for growing crops that can be consumed directly by humans in order to account for the competition for land between humans and animals. In addition, this thesis focused on European specialised dairy farming systems only. In many other dairy farming systems, however, cattle fulfil multiple functions. Due to this multi-functionality, not all systems can be analysed in the same way as the specialised systems included in this study. Neglecting the multi-functionality of cattle in, e.g. smallholder systems, will result in an overestimation of the impacts per unit of product produced. Future researches, therefore, are suggested to take into account of multi-functionality when benchmarking non-specialised dairy systems. Moreover, this thesis mainly focused on environmental performance of dairy farming systems. However, the importance of social and economic performance should not be neglected. There are potential trade-offs between environmental indicators and social or economic indicators. For example, trade-off between the GWP of a product and animal welfare indicators, or between biodiversity impacts and net farm income. Future studies, therefore are suggested to explore correlations between environmental, social and economic indicators.

Summary

Wenjuan Mu was born in Qi Xia, China in 1988. She obtained her bachelor degree cum laude in International Agribusiness & Trade at van Hall Larenstein in 2010. Her bachelor was a '2+2' programme. In the first 2 years, she studied at China Agricultural University, whereas in the last 2 years she studied at van Hall Larenstein. During her BSc, she did two internships at two Dutch companies, to gain practical experience.
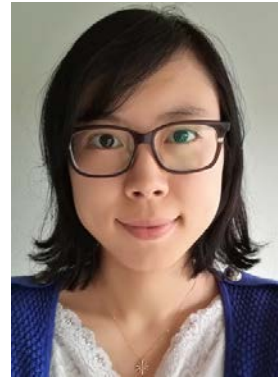
After her bachelor, she started her master study at Wageningen University. In 2012, she graduated cum laude with a master degree in Management, Economics and Consumer Studies.

In 2013, she started as a PhD student within the Animal Production Systems and Operations Research & Logistics group of Wageningen University. Her PhD was part of the IPOP project 'Mapping for Sustainable Intensification'. During her PhD, she was a member of the WIAS PhD council, where she had the opportunity to improve her organizational and communication skills and to participate actively in organising big events, such as the WIAS science day. Since July 2017, she has been working as a postdoctoral fellow within the Business Economics group of Wageningen University and Research.

About the author

Education Certificate

With the educational activities below, the PhD candidate has complied with the educational requirements set by the graduate school Wageningen Institute of Animal Science (WIAS) of Wageningen University & Research, which comprises a minimum of 30 ECTS (European Credit Transfer and Accumulation System). One ECTS equals a study load of 28 hours.

The basic package	3.0 ECTS
International conferences	3.6 ECTS
LCA Food 2014, San Francisco, USA	2014
EAAP, Warsaw, Poland	2015
Euro, Poznan, Poland	2016
Seminars and workshops	1.4 ECTS
WIAS science day, Wageningen, the Netherlands	2013-2015
Seminar nutrient management, Wageningen, the Netherlands	2014
Symposium solutions for climate change from animal production, Wageningen, the Netherlands	2014
Presentations	6.0 ECTS
Oral, WPC symposium, Wageningen, the Netherlands	2013
Poster, LCA Food conference, San Francisco, USA	2014
Poster, WIAS science day, Wageningen, the Netherlands	2014-2015
Oral, EAAP, Warsaw, Poland	2015
Oral, Euro, Poznan, Poland	2016
In-depth studies	6.4 ECTS
Sustainability analysis in food and biobased production, Wageningen, the Netherlands	2013
Theory and practice of efficiency & productivity measurement: static & dynamic analysis, Wageningen, the Netherlands	2015
International advanced course "environmental impact assessment of livestock systems, Wageningen, the Netherlands	2015
Professional skills support courses	3.9 ECTS
Techniques for writing and presenting a scientific paper, Wageningen, the Netherlands	2013
PhD competence assessment or job assessment, Wageningen, the Netherlands	2013
Entrepreneurship in and outside science, Wageningen, the Netherlands	2014
Career Perspectives, Wageningen, the Netherlands	2017
Supervision of practicals and excursions	3.0 ECTS
Sustainable development of animal systems: issues & option (APS 30306)	2015-2016

Advanced supply chain management (ORL 31306)   2016
Decision Science 1 (ORL 20306)   2016

Supervision of MSc theses   3.0 ECTS

Organisation of seminars and courses, boards and committees.   5.0 ECTS
Member of WAPS council   2015-2017
Member of WIAS Science Day committee   2016

Total completed training and supervision plan   35.2 ECTS

Acknowledgements

Doing a PhD study is a challenging task. I would like to express my sincere gratitude to all those who helped me throughout the process and made me believe that everything is possible as long as I devote myself to it. I have to say that this challenging task became nice and unforgettable, because of you.

First of all, I would like to express my truly appreciation to my supervisors Prof. dr. Imke de Boer, Prof. dr. Jacqueline Bloemhof-Ruwaard, Dr. Corina van Middelaar and Dr. Argyris Kanellopoulos. Thank you for all the guidance, advice and support during my PhD study. I am lucky to have all of you in my supervision team.

Imke was my promotor at the Animal Production Systems group. Imke, thank you for sharing your ideas and inspirations with me. You are a real scientist. You showed me that science should be dedicated and rigorous. Keeping these two words in mind certainly will benefit me substantially in my future career. Meanwhile, thank you for always being available for questions, always responding quickly to my e-mails and for providing constructive and sharp feedback to my manuscripts, even in the evenings, weekends and when you were at home for your surgery. Besides, thank you for providing me all the necessary support without doubt when I started to plan my future career, you are a mentor to both my work and life.

Jacqueline was my promotor at the Operations Research and Logistics group. Jacqueline, thank you for all your guidance throughout my entire PhD trajectory. Meanwhile, thank you very much for always trusting my research abilities and believing in me, even at moments I was doubting myself. Thank you for keeping me on track and making sure everything was finished in time. In addition, thank you very much for the talks during my pregnancy to let me know that life is not just about work, but also about family!

Corina was my co-promotor and daily supervisor at the Animal Production System group. Corina, thank you for being there during all ups and downs of my PhD trajectory. Thank you

for teaching me how to be precise in research. With your help, the quality of my work improved significantly. I would like to thank you especially for all your support during my maternity leave, and for your visits and talks after the birth of my little girl.

Argyris was my daily supervisor at Operations Research and Logistics group. Argyris, you are an excellent supervisor, I wish you would have been part of my supervision team from the very beginning of my PhD. Thank you for keeping your door always open for me. You always put the benefits of the student at the first place. Thank you for allowing me to just knock on your door whenever I ran into problems. Thank you for your patience and for your advice about defining a knowledge gap, analysing data and writing an article. To me you are not just a supervisor at work, but also a good friend with whom I can share my happiness and who provides good advices when I face difficulties.

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. I would like to acknowledge them for funding my PhD project.

I extend my sincere gratitude to all my co-authors, Jouke Oenema, Bas Engel and Evelyne Groen. It was a nice experience to cooperate with you. Thank you for all the good discussions, the papers would never be published without your contributions.

My warmest thanks to all of my colleagues in the Animal Production Systems and Operations Research and Logistics group, for their company and support. Especially, I would like to thank my officemates in both groups: Cindy, Iris, Aart, Widi, Heleen, Jochem, Aleksander, Alistair and Vitor. Thank you for all the nice events and conversations during my stay. I want to further express my thanks to Dmitry, Femke, Sonja, Marjolein, Floor, Eline and Giulia for your help and visits during my pregnancy and after the birth of Yilin. All the nice colleagues I met in both groups made my everyday PhD life much more fun.

A PhD study is a long journey I could never have accomplished without the support from my friends. I want to thank all my friends who made my PhD journey full of joy memories. Thank you, Junli Guo, Junyou Wang, Jinfeng Peng, Gongbao Wang, Hequn Li, Si Wang, Ya Wang, Jiao Long, Shumin Yu, Thomas David van der pol, Lei Mao, Man Bao, Yuan He, Wenbiao Shi, Juncai Chen, Li Meng, Fubiao Niu, Shuwen Xia, Yunmeng Zhang, Lin Ma, Xia Liu, Chunxu Song, Yue Zhao, Yuan He, Feifei Gao, Jue Wang, Xi Wan, Zhanguo Bai, Xiaoyun Bing, Zhen Liu, Qianqian Shao, Yu Hong, Tian Yu, Yun Fan, Xiangyu Hou, Kun Liu, Tingting Xiao, Tingting Zheng, Yan Wu, Lingtong Gai, Xixi Mintian, Jinghui Hao, Jingmeng Wang, Qian Liu, Xu Cheng, Guangcheng Ren and to all other friends that I may forget to mention here, please understand that my memory was kind of half gone after my pregnancy and has not fully recovered yet.

I would like to dedicate this thesis with love to my father Hongwei, my mother Wei, my husband Xuezhen and my daughter Yilin. Baba, mama, thank you for being such wonderful parents. Thank you for always being there for me whenever I need you. You are the best parents ever, you provide me enough freedom to grow and develop. Thank you for your unconditional love, your love made me strong and gave me the courage to conquer any difficulties in life. You taught me how to love and be loved, how to be a decent person. Mama, Baba, thousands of words cannot express my gratitude, everything into one sentence, I am very lucky to be your daughter and I love you forever! Meanwhile, I would like to thank my husband Xuezhen for always standing by my side. Thank you for your love, care and encouragement. I am happy that I can just be myself in front of you, you are my lover and more importantly my soul mate. Of course, I want to thank my dearest baby Yilin (Wan Baoer). Thanks for choosing me to be your mother. You are the best thing that ever happened to me, you make my life full of sunshine and happiness. Mama loves you forever!

Last but not least, I want to thank all my other family members in China, thank you for accompanying my parents when I am far away from home, thank you for all the mental support. Without all your support, I would never have managed to accomplish my PhD degree.

Acknowledgement

Colophon