Improving efficiency assessments using additive data envelopment analysis models: an application to contrasting dairy farming systems

Andreas D. Soteriades[1], Philippe Faverdin[2,3], Margaret March[1] and Alistair W. Stott[1]

[1]Scotland's Rural College, Future Farming Systems Group, Edinburgh, United Kingdom

[2]INRA, UMR 1348 PEGASE, F-35590 St-Gilles, France,

[3]Agrocampus-Ouest, UMR 1348 PEGASE, F-35000 Rennes, France

e-mail: andreas.soteriades@sruc.ac.uk

Applying holistic indicators to assess dairy farm efficiency is essential for sustainable milk production. Data Envelopment Analysis (DEA) has been instrumental for the calculation of such indicators. However, 'additive' DEA models have been rarely used in dairy research. This study presented an additive model known as slacks-based measure (SBM) of efficiency and its advantages over DEA models used in most past dairy studies. First, SBM incorporates undesirable outputs as actual outputs of the production process. Second, it identifies the main production factors causing inefficiency. Third, these factors can be 'priced' to estimate the cost of inefficiency. The value of SBM for efficiency analyses was demonstrated with a comparison of four contrasting dairy management systems in terms of technical and environmental efficiency. These systems were part of a multiple-year breeding and feeding systems experiment (two genetic lines: select vs. control; and two feeding strategies: high forage vs. low forage, where the latter involved a higher proportion of concentrated feeds) where detailed data were collected to strict protocols. The select genetic herd was more technically and environmentally efficient than the control herd, regardless of feeding strategy. However, the efficiency performance of the select herd was more volatile from year to year than that of the control herd. Overall, technical and environmental efficiency were strongly and positively correlated, suggesting that when technically efficient, the four systems were also efficient in terms of undesirable output reduction. Detailed data such as those used in this study are increasingly becoming available for commercial herds through precision farming. Therefore, the methods presented in this study are growing in importance.

Key words: efficiency indicators; slacks-based measure (SBM); undesirable outputs; savings potentials; slack shares; experimental dairy farm data

Introduction

Sustainably improving resource use efficiency to meet the nutritional demands of a growing and more affluent population is a key challenge facing humankind today (Foresight 2011). Dairy farming is required to comply with agricultural policies considering resource use efficiency as a prerequisite for sustainability (Sutton et al. 2011, van den Berg et al. 2011). Wider definitions of efficient dairy farm production are therefore essential for dairy farm sustainability and so is the translation of these definitions to improved efficiency indicators (see Callens and Tyteca 1999, de Koeijer et al. 2002). In this study, we were concerned with the calculation of more advantageous indicators for technical and environmental efficiency (from now on denoted as TE and EE respectively) of dairy farms.

Widely-used efficiency indicators evaluating dairy systems have the form of partial ratios expressing e.g. impacts per kilogram of milk and/or meat and per hectare of land use (e.g. Lovett et al. 2006, Chagunda et al. 2009, Bell et al. 2011, O'Brien et al. 2014, Ross et al. 2014). Consequently, the number of these indicators will be large if an efficiency assessment is to be comprehensive. This can complicate interpretation of the results and their communication to stakeholders (see Jollands et al. 2003). Furthermore, such partial ratios do not account for the whole range of production performance parameters (Asmild et al. 2009, Bogetoft 2012). Also, assessments of this nature tend to report indicator results averaged across all farms in the study, ignoring the potential wide range of results between different farms (see Iribarren et al. 2011).

An advantageous alternative to partial efficiency indicators is the multiple-input, multiple-output efficiency measurement method known as Data Envelopment Analysis (DEA; Charnes et al., 1978), which we employed in this study. DEA calculates single aggregated indices of efficiency for each dairy farm by assessing the whole production system and the different environmental impacts generated by the farms' given technology (Berre et al. 2013). Thus, DEA provides us with a broader view of the dairy farm efficiency problem. Consequently, numerous studies have used DEA to assess, among others, the TE and EE of dairy farms (e.g. Berre et al. 2014, Ramilan et al. 2011, Shortall and Barnes 2013, Toma et al. 2013).

Manuscript received February 2015

Current DEA dairy farm studies can be improved in three ways outlined below. The first two issues are model-related and the third is data-related. We attempted to overcome all three issues in this study. First, almost all studies have employed DEA models which assume that the reductions in inputs (e.g. feed) and undesirable outputs (e.g. greenhouse gas emissions; GHGE) that a farm should make to become efficient are proportional. The same is true for studies aiming to achieve dairy farm efficiency through an increase in desirable outputs (e.g. milk) and a decrease in undesirable outputs. On the other hand, models relaxing this proportionality assumption allow us to identify the variables (inputs and desirable/undesirable outputs) contributing the most to a farm's inefficiency (e.g. Iribarren et al. 2011). Second, most studies with undesirable outputs have used DEA models which are unable to incorporate them as actual outputs of the production process. Specifically, undesirable outputs have been modelled as inputs to be minimized (e.g. Ramilan et al. 2011, Shortall and Barnes 2013) or their reciprocal has been considered as a desirable output to be maximized (e.g. Shortall and Barnes 2013). The former practice fails to reflect the true production process (Färe and Grosskopf 2003, Kuosmanen 2005, You and Yan 2011). With the latter practice, the scale and interval of the original data are lost (You and Yan 2011). To the best of our knowledge, only a few DEA dairy studies have dealt with the issues of proportionality and inappropriate modelling of undesirable outputs (e.g. Berre et al. 2013, 2014, Iribarren et al. 2011). This can be achieved with the use of DEA models such as directional distance functions (Berre et al. 2013, 2014) or 'additive' models (Iribarren et al. 2011). In this study, we also employed additive DEA models.

The third limitation of dairy farm DEA studies is that they have often used survey data based on voluntary participation of farmers (e.g. Barnes et al. 2011, Hansson et al. 2011, Iribarren et al. 2011, Kelly et al. 2012, Shortall and Barnes 2013). This raises the question whether these data are truly random and representative (Jack 2009). Moreover, survey data are usually uncontrolled for important efficiency drivers/differentials such as management (see Cooper et al. 2007, chapter 7), genetic potential (Wall et al. 2010), feeding regime (Capper et al. 2009) and the often largely diverse climatic and bio-physical conditions under which farms operate (see Bogetoft and Otto 2011, chapter 3. Also see Barnes 2006 and Shortall and Barnes 2013). Year-to-year variation is also an important consideration in the assessment of dairy farm efficiency (e.g. Cloutier and Rowley 1993, Fogarasi and Latruffe 2009). However, farm technology can change through time in commercial herds (e.g. the introduction of a new milking parlour; Fraser and Cordina 1999) making inter-year efficiency comparisons less reliable. A solution to the aforementioned issues is to use data from experimental dairy systems which divide the herd into sub-groups ('treatments') of interest, under the same management and where more detailed data are collected to strict protocols over multiple years. An example is the multiple-year genetic line × feeding systems experiment in Dumfries, Scotland, known as the 'Langhill' experiment (Pollott and Coffey 2008).

The aims of this study were: (i) to introduce to dairy research additive DEA models that overcome the aforementioned modelling issues of proportionality and inappropriate modelling of undesirable outputs when calculating TE and EE; (ii) to demonstrate the potentials for more in-depth efficiency analyses and dairy systems comparisons using the Langhill dataset; (iii) to demonstrate how the analysis can be further informed by incorporating input price data in the study of TE; (iv) to assess whether increasing TE also increases EE; and (v) to evaluate the performance of Langhill's experimental dairy farming systems.

Material and methods

Data Envelopment Analysis

DEA was developed by Charnes et al. (1978), originating from Farrell's (1957) work. It is a non-stochastic, non-parametric technique that benchmarks different decision-making units (DMUs) performing the same task in terms of their capacity to convert inputs into outputs by using the least resources and/or producing maximum desirable output and the least undesirable output. Calculation of the aggregated DEA efficiency index does not require a priori assumptions on the importance of each variable for the DMUs' performance. This fact makes DEA a particularly attractive multiple-criteria tool. DEA constructs an efficient frontier, that is, a piece-wise linear surface over observed data points against which (the frontier) all DMUs are benchmarked. This frontier comprises of the best performers and the performance of all other DMUs is evaluated by deviations from the frontier line. This is a fundamental difference between DEA and methods such as regression as the latter reflects 'average' or 'central tendency' behaviour (Cooper et al. 2007) and is unable to provide a holistic characterization of DMUs within a multiple-objective assessment.

Dealing with proportionality in DEA

In the introduction of this study we referred to the fact that most DEA dairy farm studies have used models assuming that a DMU can become efficient by proportionally reducing its inputs and/or undesirable outputs and/or by proportionally increasing its desirable outputs. For example, assume that one uses the model of Charnes et al. (1978) to maximize efficiency by minimizing input use or by maximizing production of desirable outputs. Also assume that this model has calculated an efficiency score of 0.80 for a specific farm (with this model scores are always between zero and one). Then efficient input use for this farm is 80% of its current input use, i.e. this farm has to reduce all its inputs by 20%. Accordingly, achieving efficient desirable output levels would require this farm to increase all its desirable outputs by 20%.

There exist models not assuming that input excesses or desirable output shortfalls, called 'inefficiencies' or 'slacks' in the DEA terminology, are equal among inputs or desirable outputs. A family of such models is that of the additive models (e.g. Cooper et al. 1999, Cooper et al. 2011, Tone 2001). The term 'additive' is attributed to the fact that these models' objective functions involve summations of all input and desirable output slacks (reference to undesirable output slacks is made later in the text) in order to identify all potential sources of inefficiency. We further demonstrate this below by presenting Tone's (2001) slacks-based measure (SBM) of efficiency, variants of which were used in this study.

Suppose that there are *n* DMUs each using *m* inputs to produce *s* desirable outputs, denoted as x_{io} ($i = 1,\ldots, m$) and y_{ro} ($r = 1,\ldots, s$), all assumed positive. The SBM efficiency score of the *j*th DMU under evaluation, denoted as DMU_o, is given by the following programme (Tone 2001):

$$\rho^* = \min_{\lambda_j, s_{io}^-, s_{ro}^+} \rho = \frac{1 - \frac{1}{m}\sum_{i=1}^{m} s_{io}^-/x_{io}}{1 + \frac{1}{s}\sum_{r=1}^{s} s_{ro}^+/y_{ro}} \qquad (1)$$

subject to,

$$x_{io} = \sum_{j=1}^{n} x_{ij}\lambda_j + s_{io}^-, \qquad i = 1,\ldots,m$$

$$y_{ro} = \sum_{j=1}^{n} y_{rj}\lambda_j - s_{ro}^+, \qquad r = 1,\ldots,s$$

$$s_{io}^-, s_{ro}^+, \lambda_j \geq 0 \; (i = 1,\ldots,m, r = 1,\ldots,s, j = 1,\ldots,n),$$

where x_{io} and y_{ro} are the inputs and desirable outputs of DMU_o respectively, and s_{io}^- and s_{ro}^+ are the input and desirable output slacks of DMU_o respectively. The scalar λ_j, when greater than zero, indicates which DMUs were used as a reference by DMU_o for the calculation of ρ^*. The above programme is run *n* times, once for each DMU.

Because ρ in model 1 is minimized, its numerator is minimized and its denominator is maximized. Minimizing the numerator minimizes the negative sum of input slacks which means that it maximizes the (positive) sum of input slacks. Maximizing the denominator also maximizes the sum of desirable output slacks. Thus, model 1 calculates the maximal possible inefficiencies in inputs and desirable outputs that can occur for DMU_o relatively to the other DMUs. The value of the numerator (denominator) is at most (at least) one because the sum of input (desirable output) slacks is averaged and subtracted from (added to) unity. Thus, ρ is bounded by zero and one, with one indicating that DMU_o is efficient.

When DMU_o is efficient, all its slacks equal zero as this means that it does not need to further reduce its inputs and increase its desirable outputs to become efficient. If it is inefficient, one can identify through the slack values (which in this case are non-proportional) the inputs and desirable outputs contributing the most to its inefficiency. For an inefficient DMU any choice of input resulting in $x_{io} > \sum_j x_{ij}\lambda_j$ means that with some combination of inputs other DMUs (identified by the non-zero λ_j values) could have improved this input in amount by $s_{io}^- = x_{io} - \sum_j x_{ij}\lambda_j$ without worsening any other input or desirable output (Brockett et al. 2004). The same applies for the desirable outputs and their shortfalls $s_{ro}^+ = y_{ro} + \sum_j y_{rj}\lambda_j$.

Because the slacks are normalized by being divided by their corresponding inputs and desirable outputs, SBM is units invariant, that is, it is independent of the units in which the inputs and desirable outputs are measured, provided that these units are the same for every DMU (Cooper et al. 2007). Moreover, the normalization of slacks can be interpreted as a 'data-driven' weighting scheme. This weighting scheme is a more objective one compared to methods where the weights are (subjectively) pre-defined by the user (Cooper et al. 1999).

Modelling undesirable outputs with DEA

In the introduction we commented on the issues arising when one models undesirable outputs as inputs to be minimized or when one considers their reciprocals as desirable outputs to be maximized. Other approaches have also been suggested in the literature (see Gomes and Lins 2008, Scheel 2001), such as considering the additive inverse of an undesirable output as a desirable output and then adding to this a positive scalar large enough to convert it to a positive value. These methods still fail to reflect the true production process and also cannot be applied to all DEA models. On the other hand, undesirable outputs can be modelled as such in an additive manner with the use of SBM models. Specifically, Tone's (2001) SBM (model 1) has been extended to the Undesirable Output Model (Cooper et al. 2007) presented below.

Suppose that, in addition to the inputs and desirable outputs defined above, the n DMUs also produce k undesirable outputs, denoted as z_d ($d = 1,\ldots, k$) respectively, assumed positive. The Undesirable Output Model is the following programme:

$$\rho^{u*} = \min_{\lambda_j, s_{io}^-, s_{ro}^+, s_{do}^u} \rho^u = \frac{1 - \frac{1}{m}\sum_{i=1}^{m} s_{io}^- / x_{io}}{1 + \frac{1}{s+k}\left(\sum_{r=1}^{s} s_{ro}^+ / y_{ro} + \sum_{d=1}^{k} s_{do}^u / z_{do}\right)} \tag{2}$$

subject to

$$x_{io} = \sum_{j=1}^{n} x_{ij}\lambda_j + s_{io}^-, \qquad i = 1, \ldots, m$$

$$y_{ro} = \sum_{j=1}^{n} y_{rj}\lambda_j - s_{ro}^+, \qquad r = 1, \ldots, s$$

$$z_{do} = \sum_{j=1}^{n} z_{dj}\lambda_j + s_{do}^u, \qquad d = 1, \ldots, k$$

$$s_{io}^-, s_{ro}^+, s_{do}^u, \lambda_j \geq 0 \; (i = 1, \ldots, m, r = 1, \ldots, s, d = 1, \ldots, k, j = 1, \ldots, n),$$

where z_{do} and s_{do}^u are the undesirable outputs and their slacks for DMU$_o$ respectively. Importantly, in model 2 the undesirable outputs are positioned in the output set and neither transformation of their values (e.g. reciprocal, additive inverse, etc.) nor position change from output to input are required (Cooper et al. 2007). As with model 1, model 2 is run n times, $\rho^{u}*$ takes values between zero and one and an efficient DMU has all its slacks equal to zero.

An important aspect of undesirable output modelling is that of disposability, which refers to the impact that undesirable output reduction can have on inputs and desirable outputs. Specifically, undesirable outputs are weakly disposable when they cannot be reduced without increasing inputs or reducing desirable outputs. On the other hand, strongly disposable undesirable outputs can be reduced at no cost. (The readers are referred to Färe et al. 1989 as a starting point to disposability. Also, see Yang and Pollitt 2010 for a comprehensive coverage of references and for example models.)

In dairy studies with DEA, undesirable outputs have been modelled as both weakly (e.g. Berre et al. 2013, 2014, Ramilan et al. 2011, Toma et al. 2013) and strongly disposable (e.g. Iribarren et al. 2011, Shortall and Barnes 2013). Shortall and Barnes (2013), whose study's undesirable outputs were GHGE, followed the argument of de Koeijer et al. (2002) that environmental impacts generated by non-point source pollutants must be addressed through more efficient input use rather than desirable output reduction. In this study, we adopted this logic to model GHGE, the undesirable output of our study (see Data sub-section below). However, it should be pointed out that weak disposability of undesirable outputs can be added to model 2 if necessary, see Bremberger et al. (2015).

Technical and environmental efficiencies with SBM

TE can be defined as a DMU's ability to minimize its inputs given its current (desirable and undesirable) output production. TE can also be defined as a DMU's ability to maximize desirable output given its current input use and production of undesirable outputs. The former definition of TE refers to an input-oriented model while the latter to an output-oriented one. To demonstrate the TE model below, we used input orientation. Model 1 can be easily modified to an input-oriented TE measure as follows (Cooper et al. 2007):

$$\mathrm{TE} = \min_{\lambda_j, s_{io}^-} 1 - \frac{1}{m}\sum_{i=1}^{m} s_{io}^- / x_{io} \tag{3}$$

subject to

$$x_{io} = \sum_{j=1}^{n} x_{ij}\lambda_j + s_{io}^-, \qquad i = 1, \dots, m$$

$$y_{ro} \leq \sum_{j=1}^{n} y_{rj}\lambda_j\,, \qquad r = 1, \dots, s$$

$$z_{do} \geq \sum_{j=1} z_{dj}\lambda_j\,, \qquad d = 1, \dots, k$$

$$s_{io}^-, \lambda_j \geq 0\ (i = 1, \dots, m, j = 1, \dots, n),$$

where the third constraint is added because undesirable outputs were also present in our production technology (see Ramilan et al. 2011). It should be noted that model 3 is a simple linear programme. Model 3 has also been used in the study of Iribarren et al. (2011).

EE is defined as a DMU's ability to minimize undesirable outputs for the given input and desirable output levels (see Ramilan et al. 2011, Shortall and Barnes 2013). This leads to an undesirable output-oriented model. Model 2 can be modified so as to agree with this definition of EE as follows:

$$\mathrm{EE} = \min_{\lambda_j, s_{do}^u} \frac{1}{1 + \frac{1}{k}\sum_{d=1}^{k} s_{do}^u / z_{do}} \tag{4}$$

subject to

$$x_{io} \geq \sum_{j=1}^{n} x_{ij}\lambda_j\,, \qquad i = 1, \dots, m$$

$$y_{ro} \leq \sum_{j=1}^{n} y_{rj}\lambda_j\,, \qquad r = 1, \dots, s$$

$$z_{do} = \sum_{j=1}^{n} z_{dj}\lambda_j + s_{do}^u, \qquad d = 1, \dots, k$$

$$s_{do}^u, \lambda_j \geq 0\ (d = 1, \dots, k, j = 1, \dots, n).$$

Model 4 can be converted to a simple linear programme by maximizing the objective function's denominator instead and then calculating its reciprocal to obtain the EE score.

Once the models for TE and EE are run, the optimal input and undesirable output slacks for DMU_o can be used to examine variable-specific patterns, i.e. input and undesirable output savings potentials, by calculating the ratio of each slack over its corresponding input or undesirable output.

Seeking cost-effective reductions in inputs

One can assign input price data to the optimal slacks s_{io}^{-*}(the asterisk denotes optimality) calculated by model 3 in order to determine the proportion of their cost to the total input cost. We refer to this proportion as 'slack share' (Tsutsui and Goto 2009). Specifically, let us define p_{io} ($i = 1,..., m$) as the price of input x_{io} for DMU$_o$. Then the slack share of slack s_{io}^{-*} for DMU$_o$ is given by the following ratio:

$$Slack\ share_{io} = \ (p_{io}s_{io}^{-*})/\sum_{i=1}^{m} p_{io}x_{io} \ \ (i = 1, \dots, m) \tag{5}$$

Comparing a DMU's input savings potentials with the slack shares reveals whether there exist both technical and economic incentives for this DMU to reduce its inputs to the technically efficient levels. For example, consider that DMU$_o$ has a large slack share for a specific input. Then, reducing this input by its slack would be effective for reducing the overall cost of DMU$_o$. On the other hand, reducing an input with a large savings potential and a small slack share to its technically efficient levels would save only a small amount of the overall cost (see Tsutsui and Goto 2009).

Data

Our study used detailed records from Scotland's Rural College's (SRUC) Langhill dairy systems study (Veerkamp et al. 1994, Pollott and Coffey 2008), a long-term breeding and feeding systems experiment. The production systems within the herd represented a range of dairy systems that may be found commercially. The herd consisted of genetic lines selected for kilograms of milk fat plus protein or selected to remain close to the average genetic merit for milk fat plus protein production for all animals evaluated in the UK each year. The advantage of this dataset is that it provided information that is not routinely available on commercial farms and allowed for efficiency comparisons between clearly-defined systems; and between years. In the next sub-sub-section we briefly describe the experiment, based on the studies of Bell et al. (2011), Chagunda et al. (2009), Pryce et al. (1999) and Ross et al. (2014). The second sub-sub-section is devoted to the description of the DEA variables and their derivation. These DEA variables were derived by Toma et al. (2013) for the purpose of their own DEA exercise so readers are referred to their study for further details. The third sub-sub-section describes the input price data used for the calculation of input slack shares.

Langhill dairy systems experiment

The data used in this study covered the period 2004–2010, during which time the experiment's protocol remained unchanged. The herd was divided into four distinct systems defined by two different genetic merits fed on two different diets. The number of cows was kept at approximately 50 per system. The high forage diet aimed at providing 70–75% home-grown forage in the dry matter (DM), complemented by bought-in concentrates and summer grazing, typically from March to November. The low forage herds were housed all-year-round, and their diet consisted of about 45% home-grown forages and 55% bought-in concentrates. Within each diet, cows were either of average UK genetic merit for milk fat and protein production (control cows) or represented the top 5% of UK genetic merit (select cows). Thus, there were four distinct systems, namely high-forage control (HFC), high-forage select (HFS), low-forage control (LFC) and low-forage select (LFS).

Cows remained in the herd for at least three lactations unless culling was necessary due to reduced cow welfare. Cows of greater than three lactations could be retained in the herd until a replacement heifer of suitable genetic merit was available. Cows were milked three times a day and milk yield per cow was automatically recorded, while fat and protein concentrations in milk were obtained from weekly cow-specific samples. Live-weights were recorded after every milking for milking cows and weekly for dry cows and replacement animals. Feed intakes of individual milking cows were recorded using automated HOKO feed measurement gates (Insentec BV, Marknesse, The Netherlands).

Other data recorded in the Langhill dataset included annual home-grown forage yields and on-farm fertilizer applications, land use and fuel use. Annual crop yields and hectares required were obtained directly from the database or from farm records. The Langhill database holds information on each farm field and all activities that are carried out such as sowing, fertilizing, and also the number of trailer loads harvested. Types of fertilizer used and application rates are also routinely recorded.

DEA variables and how they were derived

In this sub-sub-section we outline the DEA variables that we used for our exercise. Then, we briefly describe the process followed by Toma et al. (2013) in order to derive these variables. For further details refer to their study.

We used the following DEA input, output and undesirable output data per system and per year for the calculation of TE and EE:

- Inputs: home-grown feed (forage and grazed grass; t [tonnes] DM); purchased feed (concentrates; t DM); land use (ha); and nitrogen (N) fertilizer use (t N). It should be noted that land use and N fertilizer use data concerned only on-farm activities, i.e. land use and N fertilizer use embedded in purchased feed were not accounted for. Other inputs such as labour and capital were not available in the database and thus were not included in the input dataset. Replacements were also not included because replacement rates between systems were similar (the experiment's protocol required that all cows remain in the system for maximum three lactations only).

- Output: energy-corrected milk (t; see Sjaunja et al. 1990).

- Undesirable output: GHGE (t CO_2-eq. [CO_2-equivalents]).

Statistics for the data above per system and per year are presented in the Appendix.

We assumed that the systems' technology did not change from year to year and so we considered each system as a different DMU for each of the seven years of the experiment. This resulted in a total of (seven years) × (four systems) = 28 DMUs. This assumption was based on the fact that the experiment's protocol and management practices remained unchanged during the years 2004–2010 and so did the systems' production technology.

The DEA variables were derived as follows. Data were extracted from the Langhill database for each cow and aggregated annually at the four system levels for each of the seven years. Data relating to milk yield, fertilizer application, fuel use, feed intake, land use and diet were extracted directly from the database and data for herd dynamics and young stock were taken from an annual inventory of the systems. Daily milk yields were summed by system and fat and protein concentrations from each of the three daily milking times were sampled and analysed once per week and averaged.

Aggregated annual system data were then used to calculate total GHGE, expressed in t CO_2-eq. The PAS2050 accredited SAC Carbon Calculator v3.11 (SAC 2011) estimated GHGE attributed to each of the four groups. The Carbon Calculator applies IPCC (2006) Tier 2 methodologies, equations, and emission factors (outlined in Table 1 in Toma et al. 2013) and requires detailed information regarding farm inputs and outputs related to livestock, land and crops, purchased feeds and energy use. Herd dynamic inputs were based on an annual reconciliation of all ages of livestock accounting for sales, purchases and cow culling. Direct and indirect CO_2 emissions were calculated by allocating land hectares and fertilizer applications on a per-crop and fertilizer-type basis. Electricity use was not available in the database and was estimated from milk yield (Sheane et al. 2010). Due to limited use or data unavailability, sprays of pesticides, fungicides, herbicides and carbon sequestration were not accounted for in the carbon foot-printing exercise. Also, emissions from meat production were not included because replacement rates were similar between systems, the destinations of culled cows varied, and the emissions from calves leaving the system for beef would be attributable to the farms raising them.

Input price data

By comparing the TE input savings potentials with their slack shares (derived from equation 5) showed whether or not it was cost-effective to reduce current input use (see sub-sub-section 'Seeking cost-effective reductions in inputs' above). For that purpose we used input price data for the financial year April 2010–March 2011 (Table 1). Therefore, the comparison between input savings potentials and slack shares was restricted to systems HFC, HFS, LFC and LFS for the year 2010–2011 only so as to demonstrate equation 5. The price data in Table 1 were obtained from Langhill's accounting data, the Farm Management Handbook (SAC Consulting 2010) and from DairyCo's online price data (http://www.dairyco.org.uk/market-information/farm-expenses/#.VGsUj2d5erE).

Table 1. Input price data for each system for the financial year April 2010–March 2011

	HFC	HFS	LFC	LFS
Home-grown feed [£ (t DM)$^{-1}$]	37.55	38.16	43.72	44.57
Purchased feed [£ (t DM)$^{-1}$]	196.95	198.87	205.01	209.02
Land use (£ ha^{-1})	180.00	180.00	180.00	180.00
N fertilizer [£ (t N)$^{-1}$]	756.52	756.52	756.52	756.52

HFC = high-forage control; HFS = high-forage select; LFC = low-forage control; LFS = low-forage select; DM = Dry Matter; N = nitrogen; t = tonnes

On the calculation of indicators for TE and EE

The data for land use, N fertilizer use and home-grown feed were interrelated as, for example, minimizing home-grown feed involves reduced fertilizer requirements (Iribarren et al. 2011). Therefore, including land use, N fertilizer use and home-grown feed in the same input set implies, to a large extent, double-counting of these three inputs. This can result in unreliable slack values for these three inputs. Consequently, we ran one DEA exercise with home-grown feed and purchased feed as the only inputs; and one where purchased feed was kept as an input but home-grown feed was replaced by two 'proxies', that is, land use and N fertilizer use (e.g. Kelly et al. 2012). We denote TE and EE for the first DEA run as TE^F and EE^F (F for 'feeds') respectively. For the second run, TE and EE are denoted as TE^{LN} and EE^{LN} (L for 'land use' and N for 'N fertilizer') respectively. We calculated the correlations between TE^F and TE^{LN} and between EE^F and EE^{LN} to test the degree to which these indicators were interchangeable with each other. The correlations were calculated with the non-parametric Spearman's rank correlation coefficient. This coefficient quantifies monotone dependence between two variables by ranking their values (Panik 2005).

All DEA and statistical calculations were run in the programming language R (R Development Core Team, 2014).

Application to the Langhill data

Correlations between indicators for TE and EE

Indicator interchangeability

The Spearman's rank correlation between TE^F and TE^{LN} was strong (0.78), while that between EE^F and EE^{LN} was near-perfect (0.99). These findings suggested that our four indicators for TE and EE were interchangeable. Thus, all four indicators were used in the analysis below.

Synergy between TE and EE

The (Spearman's rank) correlations between TE^F and EE^F; and between TE^{LN} and EE^{LN} were strong (0.65 and 0.85 respectively). This result was in line with the findings of Shortall and Barnes (2013) that more technically efficient dairy farms are also more environmentally efficient in terms of GHGE.

Efficiency scores per efficiency type and system

The efficiency scores per system for TE and EE are summarized in Table 2. For TE^F the highest-to-lowest efficiency systems were, both in terms of mean and median, HFS > LFS > HFC > LFC. For EE^F, TE^{LN} and EE^{LN} the highest-to-lowest efficiency systems were, both in terms of mean and median, LFS > HFS > LFC > HFC. Therefore, systems LFS and HFS, both systems with cows of high (select) genetic potential for milk production, were the best performers for TE and EE. This suggested that select cows might be better able to achieve higher efficiency performance, regardless of whether on a low-forage or a high-forage diet. Our findings for EE agreed with those of Toma et al. (2013). Our additional contribution to their findings was the calculation of TE which shows, for the case of TE^F, that HFS systems could outperform LFS systems. However, the efficiency scores' standard deviations for LFS and HFS systems indicated a higher variation of TE and EE scores between years. This was also true for the HFC system for TE^{LN}. On the other hand, the LFC system was the least variable for both TE and EE at the expense of lower minimum and maximum efficiency scores (Table 2).

Table 2. DEA efficiency scores for TE and EE per system averaged across the years 2004–2010

	TE^{F}				TE^{LN}			
	HFC	HFS	LFC	LFS	HFC	HFS	LFC	LFS
Mean	0.79	0.92	0.78	0.90	0.72	0.89	0.83	0.95
SD	0.05	0.06	0.03	0.08	0.08	0.09	0.04	0.07
Median	0.80	0.91	0.79	0.86	0.76	0.90	0.84	1.00
Min	0.72	0.84	0.74	0.80	0.63	0.76	0.77	0.81
Max	0.84	1.00	0.82	1.00	0.82	1.00	0.88	1.00
	EE^{F}				EE^{LN}			
	HFC	HFS	LFC	LFS	HFC	HFS	LFC	LFS
Mean	0.76	0.90	0.84	0.93	0.76	0.90	0.84	0.96
SD	0.03	0.07	0.03	0.06	0.03	0.07	0.03	0.05
Median	0.77	0.88	0.85	0.94	0.77	0.88	0.85	1.00
Min	0.72	0.83	0.79	0.88	0.72	0.83	0.79	0.89
Max	0.79	1.00	0.88	1.00	0.79	1.00	0.88	1.00

TE = technical efficiency; EE = environmental efficiency; TE^{F} and EE^{F} = TE and EE, respectively, with home-grown feed and purchased feed as inputs; TE^{LN} and EE^{LN} = TE and EE, respectively, with land use, N fertilizer use and purchased feed as inputs; HFC = high-forage control; HFS = high-forage select; LFC = low-forage control; LFS = low-forage select; SD = standard deviation

Savings potentials

Using models 3 and 4 for the calculation of TE and EE respectively facilitated the identification of those inputs and undesirable outputs contributing the most to the DMUs' inefficiency. This allowed for the identification of specific aspects in which the four systems differed. As discussed above, this can be done by looking at the input and undesirable output savings potentials. These savings potentials are summarised per efficiency type and system in Table 3.

The savings potentials of system HFC for home-grown feed, land use and N fertilizer use were by far the largest (Table 3). Also, the same savings potentials were notably large for the system LFC, and they were the smallest for LFS. Two main conclusions were drawn from these results. First, the results confirmed our findings above, i.e. systems HFC and LFC, comprising of control cows, could not compete in terms of resource use efficiency with systems HFS and LFS, comprising of select cows (see Veerkamp et al. 1994). Second, they reflected the high dependency of HFC on home-grown feed and associated land use and N fertilizer use. These two conclusions were justified by looking at the same savings potentials for HFS: they were also large, but not always larger than those of LFC. In terms of purchased feed savings potentials, systems LFC and LFS were comparatively disadvantaged as a larger part of their diet depended on purchased feed. It should be noted though that the difference in purchased feed savings potentials between LFC and LFS was quite large, with the latter system performing better.

Table 3. Input and undesirable output savings potentials per system and efficiency type.

	TE^{F}		TE^{LN}			EE^{F}	EE^{LN}
	Home-grown feed	Purchased feed	Land use	N fertilizer	Purchased feed	GHGE	
HFC	42.7	0.9	39.3	48.2	0.9	32.5	32.5
HFS	16.3	0.0	11.6	21	0.0	11.3	11.2
LFC	20.5	23.2	12.4	20.1	21.8	19.8	19.4
LFS	9.8	10.5	4.1	6.3	5.9	7.0	4.0

TE^{F} and EE^{F} = technical and environmental efficiency, respectively, with home-grown feed and purchased feed as inputs; TE^{LN} and EE^{LN} = technical and environmental efficiency, respectively, with land use, N fertilizer use and purchased feed as inputs; HFC = high-forage control; HFS = high-forage select; LFC = low-forage control; LFS = low-forage select; GHGE = greenhouse gas emissions

The lowest to highest ranking of the four systems for GHGE savings potentials was LFS > HFS > LFC > HFC (Table 3). This was an important finding in that it demonstrated the advantage of aggregated DEA indicators over partial ratios. Indeed, Toma et al. (2013) expressed the four systems' GHGE per kg of energy-corrected milk and found that the systems ranked as LFS > LFC > HFS > HFC. This partial ratio of efficiency ignored other production factors (feed, land, N fertilizer), leading to the conclusion that LFC systems are more environmentally efficient than HFS.

Our finding also demonstrated the usefulness of non-proportional slacks for ranking DMUs in terms of input and undesirable output-specific performance.

Savings potentials and slack shares for the year 2010–2011

In this sub-section we demonstrate the benefit of introducing input price data to the analysis of TE. We calculated the four systems' slack shares for the year 2010–2011 (the last year of the experiment) with equation 5 and compared these to the systems' input savings potentials for the same year (Table 4).

The savings potentials of systems HFS and LFS were zero for the year 2010–2011 so we focused the analysis on HFC and LFC systems (Table 4). Notably, the savings potentials of system HFC for home-grown feed, land use and N fertilizer use were large (between 49.3% and 54%). However, the corresponding slack shares ranged between 2.3% and 7.9%. The same was true for LFC but to a lesser extent. On the other hand, both the savings potential and slack share of purchased feed were high for LFC.

Table 4. Input savings potentials and slack shares per system for the year 2010–2011

	TE^F				TE^{LN}					
	Home-grown feed		Purchased feed		Land use		N fertilizer		Purchased feed	
	Savings potential	Slack share	Savings potential	Slack share	Savings potential	Slack share	Savings potential	Slack share	Savings potential	Slack share
HFC	50.5	7.9	2.6	2.2	54	6.4	49.3	2.3	2.6	2.2
HFS	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
LFC	17.7	1.6	21.6	19.7	18.8	1.1	20.1	0.5	21.6	19.9
LFS	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0

TE^F and EE^F = technical and environmental efficiency, respectively, with home-grown feed and purchased feed as inputs; TE^{LN} and EE^{LN} = technical and environmental efficiency, respectively, with land use, N fertilizer use and purchased feed as inputs; HFC = high-forage control; HFS = high-forage select; LFC = low-forage control; LFS = low-forage select

These results were particularly interesting for home-grown feed, purchased feed and N fertilizer use, for the following reason. Comparing the slack shares with the savings potentials helped prioritize the reduction of those inputs that simultaneously resulted in increased TE and economic savings. For example, LFC clearly had to prioritize the reduction of purchased feed over home-grown feed and N fertilizer use (Table 4). Such an analysis with SBM would be greatly beneficial for commercial farmers, especially when combined with further economic analyses on the slacks. For instance, in the study of Iribarren et al. (2011), the slacks derived from model 3 were 'priced' to show that Galician dairy farms could achieve significant economic savings.

Conclusions

In this study we demonstrated the advantages of additive SBM models for the calculation and analysis of the efficiency of dairy systems. The SBM models aggregate the non-proportional input and (desirable and undesirable) output slacks into single efficiency scores without requiring a priori weighting of the inputs and outputs. The efficiency scores are dimensionless and bounded between 0 and 1, thus allowing for the comparison of different DMUs in terms of efficiency. The slacks can be used for further efficiency analyses, for example to determine the contribution of each DEA variable to each DMU's inefficiency. Also, the input slacks can be 'priced' so as to determine the cost of inefficient input use. Moreover, with SBM models undesirable outputs are positioned in the output set. Thus, neither transformation of their values (e.g. reciprocal, additive inverse, etc.) nor position change from output to input are required for the calculation of EE.

In this study, the SBM models were applied to compare the four Langhill systems, namely, HFC, HFS, LFC and LFS. The Langhill dataset was particularly advantageous for systems comparisons because it consisted of detailed data collected to strict protocols over multiple years. Importantly, efficiency drivers/differentials such as farm management and bio-physical conditions did not apply as all four systems were in the same farm. Our main conclusions are listed below:

The strong and positive correlation between TE and EE suggested that when technically efficient, the four systems were also efficient in terms of GHGE reduction.

The better efficiency performance for TE and EE of LFS and HFS systems compared to that of HFC and LFC systems showed that select animals could outperform control animals in terms of efficiency regardless of feeding strategy. However, HFS and LFS systems were more volatile from year to year for TE and EE.

These results require further testing with larger datasets. They demonstrated, however, the value of SBM for this type of analyses. Because of recent advancements in precision agriculture for commercial herds (e.g. HM Government 2013), the use of more advantageous efficiency models such as SBM grows in importance.

Acknowledgements

Thank you to all Crichton Royal Farm staff and to all SAC Consulting staff who helped with the data; as well as to Dave Roberts and Jennifer Flockhart at Scotland's Rural College for the same reason. Also, thank you to three anonymous reviewers who helped significantly improve earlier versions of this manuscript. This project is jointly funded by Scotland's Rural College (SRUC) and the French National Institute for Agricultural Research (INRA). It is also supported by, and benefits from, Scottish Government's Strategic Research Programmes.

References

Asmild, M., Holvad, T., Hougaard, J.L. & Kronborg, D. 2009. Railway reforms: do they influence operating efficiency? *Transportation* 36: 617–638.

Barnes, A.P. 2006. Does multi-functionality affect technical efficiency? A non-parametric analysis of the Scottish dairy industry. *Journal of Environmental Management* 80: 287–294.

Barnes, A.P., Rutherford, K.M.D., Langford, F.M. & Haskell, M.J. 2011. The effect of lameness prevalence on technical efficiency at the dairy farm level: an adjusted data envelopment analysis approach. *Journal of Dairy Science* 94: 5449–5457.

Bell, M.J., Wall, E., Russell, G., Simm, G. & Stott, A.W. 2011. The effect of improving cow productivity, fertility, and longevity on the global warming potential of dairy systems. *Journal of Dairy Science* 94: 3662–3678.

Berre, D., Blancad, S.B., Boussemart, J.-P. & Leleu, H. 2014. Finding the right compromise between productivity and environmental efficiency on high input tropical dairy farms: a case study. *Journal of Environmental Management* 146: 235–244.

Berre, D., Boussemart, J.-P., Leleu, H. & Tillard, E. 2013. Economic value of greenhouse gases and nitrogen surpluses: society vs farmers' valuation. *European Journal of Operational Research* 226: 325–331.

Bogetoft, P. 2012. *Performance benchmarking. Measuring and managing performance*. New York: Springer Science+Business Media. 255 p.

Bogetoft, P. & Otto, L. 2011. *Benchmarking with DEA, SFA, and R*. New York: Springer. p. 57–80.

Bremberger, C., Bremberger, F., Luptacik, M. & Schmitt, S. 2015. Regulatory impact of environmental standards on the eco-efficiency of firms. *Journal of the Operational Research Society* 66: 421–433.

Brockett, P.L., Cooper, W.W., Golden, L.L., Rousseau, J.J. & Wang, Y. 2004. Evaluating solvency versus efficiency performance and different forms of organization and marketing in US property-liability insurance companies. *European Journal of Operational Research* 154: 492–514.

Callens, I. & Tyteca, D. 1999. Towards indicators of sustainable development for firms: a productive efficiency perspective. *Ecological Economics* 28: 41–53.

Capper, J.L., Cady, R.A. & Bauman, D.E. 2009. The environmental impact of dairy production: 1944 compared with 2007. *Journal of Animal Science* 87: 2160–2167.

Chagunda, M.G.G., Römer, D.A.M. & Roberts, D.J. 2009. Effect of genotype and feeding regime on enteric methane, non-milk nitrogen and performance of dairy cows during the winter feeding period. *Livestock Science* 122: 323–332.

Charnes, A., Cooper, W.W. & Rhodes, E. 1978. Measuring the efficiency of decision making units. *European Journal of Operational Research* 2: 429–444.

Cloutier, L.M. & Rowley, R. 1993. Relative technical efficiency: data envelopment analysis and Quebec's dairy farms. *Canadian Journal of Agricultural Economics* 41: 169–176.

Cooper, W.W., Park, K.S. & Pastor, J.T. 1999. RAM: A range adjusted measure of inefficiency for use with additive models, and relations to other models and measures in DEA. *Journal of Productivity Analysis* 11: 5–42.

Cooper, W.W., Pastor, J.T., Borras, F., Aparicio, J. & Pastor, D 2011. BAM: a bounded adjusted measure of efficiency for use with bounded additive models. *Journal of Productivity Analysis* 35: 85–94.

Cooper, W.W., Seiford, L.M. & Tone, K. 2007. *Data Envelopment Analysis: A comprehensive text with models, applications, references and DEA-solver software*. Springer Science+Business Media, LLC. 490 p.

de Koeijer, T.J., Wossink, G.A.A., Struik, P.C. & Renkema, J.A. 2002. Measuring agricultural sustainability in terms of efficiency: the case of Dutch sugar beet growers. *Journal of Environmental Management* 66: 9–17.

Färe, R. & Grosskopf, S. 2003. Nonparametric productivity analysis with undesirable outputs: comment. *American Journal of Agricultural Economics* 85: 1070–1074.

Färe, R. & Grosskopf, S., Lovell C.A.K. & Pasurka, C. 1989. Multilateral productivity comparisons when some outputs are undesirable: a nonparametric approach. *The Review of Economics and Statistics* 71: 90–98.

Farrell, M.J. 1957. The measurement of productive efficiency. *Journal of the Royal Statistical Society: Series A* 120: 253–290.

Fogarasi, J. & Latruffe, L. 2009. Technical efficiency in dairy farming: a comparison of France and Hungary in 2001-2006. *Studies in Agricultural Economics* 110: 75–84.

Foresight 2011. *The future of food and farming: final project report*. London: The Government Office for Science. 208 p. https://www.gov.uk/government/uploads/system/uploads/attachment_data/file/288329/11-546-future-of-food-and-farming-report.pdf. Accessed 28 September 2015.

Fraser, I. & Cordina, D. 1999. An application of data envelopment analysis to irrigated dairy farms in Northern Victoria, Australia. *Agricultural Systems* 59: 267–282.

Gomes, E.G. & Lins, M.P.E. 2008. Modelling undesirable outputs with zero sum gains data envelopment analysis models. *Journal of the Operational Research Society* 59: 616–623.

Hansson, H., Szczensa-Rundberg, M. & Nielsen, C. 2011. Which preventive measures against mastitis can increase the technical efficiency of dairy farms? *Animal* 5: 632–640.

HM Government 2013. *A UK strategy for agricultural technologies*. London: Her Majesty's Government. 50 p. https://www.gov.uk/government/uploads/system/uploads/attachment_data/file/227259/9643-BIS-UK_Agri_Tech_Strategy_Accessible.pdf. Accessed 28 September 2015.

IPCC 2006. IPCC Guidelines for National Greenhouse Gas Inventories, Prepared by the National Greenhouse Gas Inventories Programme. In: Eggleston H.S., Buendia L., Miwa K., Ngara T. and Tanabe K. (eds). Hayama: IGES. http://www.ipcc-nggip.iges.or.jp/public/2006gl/index.html. Accessed 28 September 2015.

Iribarren, D., Hospido, A., Moreira, M.T. & Feijoo, G. 2011. Benchmarking environmental and operational parameters through eco-efficiency criteria for dairy farms. *Science of the Total Environment* 409: 1786–1798.

Jack, L. 2009. *Benchmarking in food and farming: creating sustainable change*. Farnham: Gower Publishing Limited. 148 p.

Jollands N., Lermit J. & Patterson M. 2003. The usefulness of aggregate indicators in policy-making and evaluation: a decision with application to eco-efficiency indicators in New-Zealand. The International Society for Ecological Economics. http://www.isecoeco.org/pdf/jollands.pdf. Accessed 28 September 2015.

Kelly, E., Shalloo, L., Geary, U., Kinsella, A., Thorne, F. & Wallace, M. 2012. The associations of management and demographic factors with technical, allocative and economic efficiency of Irish dairy farms. *Journal of Agricultural Science* 150: 738–754.

Kuosmanen, T. 2005. Weak disposability in nonparametric production analysis with undesirable outputs. *American Journal of Agricultural Economics* 87: 1077–1082.

Lovett, D.K., Shalloo, L., Dillon, P. & O'Mara, F.P. 2006. A systems approach to quantify greenhouse gas fluxes from pastoral dairy production as affected by management regime. *Agricultural Systems* 88: 156–179.

O'Brien, D., Capper, J.L., Garnsworthy, P.C., Grainger, C. & Shalloo, L. 2014. A case study of the carbon footprint of milk from high-performing confinement and grass-based dairy farms. *Journal of Dairy Science* 97: 1835–1851.

Panik 2005. *Advanced statistics from an elementary point of view*. Amsterdam: Elsevier. 802 p.

Pollott, G.E. & Coffey, M.P. 2008. The effect of genetic merit and production system on dairy cow fertility, measured using progesterone profiles and on-farm recording. *Journal of Dairy Science* 91: 3649–3660.

Pryce, J.E., Nielsen, B.L., Veerkamp, R.F. & Simm, G. 1999. Genotype and feeding system effects and interactions for health and fertility traits in dairy cattle. *Livestock Production Science* 57: 193–201.

R Development Core Team 2014. R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. http://www.R-project.org/

Ramilan, T., Scrimgeour, F. & March, D. 2011. Analysis of environmental and economic efficiency using a farm population micro-simulation model. *Mathematics and Computers in Simulation* 81: 1344–1352.

Ross, S.A., Chagunda, M.G.G., Topp, C.F.E. & Ennos, R. 2014. Effect of cattle genotype and feeding regime on greenhouse gas emissions intensity in high producing dairy cows. *Livestock Science* 170: 158–171.

SAC 2011. *Carbon calculator v3.11*. Penicuik : SAC Consulting.

SAC Consulting 2010. *The farm management handbook 2010/11. The UK reference for farm business management*. Edinburgh: SAC Consulting. 342 p.

Scheel, H. 2001. Undesirable outputs in efficiency evaluations. *European Journal of Operational Research* 132: 400–410.

Sheane, R., Lewis, K., Hall, P., Holmes-Ling, P., Kerr, A., Stewart, K. & Webb, D. 2011. *Identifying opportunities to reduce the carbon footprint associated with the Scottish dairy supply chain —Methodology report*. Edinburgh: The Scottish Government. 54 p. http://www.gov.scot/Resource/Doc/342351/0113918.pdf. Accessed 28 September 2011.

Shortall, O.K. & Barnes, A.P. 2013. Greenhouse gas emissions and the technical efficiency of dairy farmers. *Ecological Indicators* 29: 478–488.

Sjaunja, L.O., Baevre, L., Junkkarinene, L., Pedersen, J. & Setala, J. 1990. A Nordic proposal for an energy corrected milk (ECM) formula. In: *Proceedings of the 27th Session of the International Commission for Breeding and Productivity of Milk Animals*, Paris, France. p. 156–157.

Sutton, M.A., Howard, C.M., Erisman, J.W., Billen, G., Bleeker, A., Grennfelt, P., van Grinsven, H. & Grizzetti, B. 2011. *The European Nitrogen Assessment*. New York: Cambridge University Press. 612 p.

Toma, L., March, M., Stott, A.W. & Roberts, D.J. 2013. Environmental performance of dairy systems: a productive efficiency approach. *Journal of Dairy Science* 96: 7014–7031.

Tone, K. 2001. A slacks-based measure of efficiency in data envelopment analysis. *European Journal of Operational Research* 130: 498–509.

Tsutsui, M. & Goto, M. 2009. A multi-division efficiency evaluation of U.S. electric power companies using a weighted slacks-based measure. *Socio-Economic Planning Sciences* 43: 201–208.

van den Berg, M., Bakkes, J., Bouwman, L., Jeuken, M., Kram, T., Neumann, K., van Vuuren, D.P. & Wilting, H. 2011. *EU resource efficiency perspectives in a global context*. The Hague: PBL Netherlands Environmental Assessment Agency. 107 p. http://www.pbl.nl/sites/default/files/cms/publicaties/PBL-EU-Resource-Efficiency-Perspectives_web.pdf. Accessed 28 September 2015.

Veerkamp, R.F., Simm, G. & Oldham, J.D. 1994. Effects of interaction between genotype and feeding system on milk production, feed intake, efficiency and body tissue mobilization in dairy cows. *Livestock Production Science* 39: 229–241.

Wall, E., Simm, G. & Moran, D. 2010. Developing breeding schemes to assist mitigation of greenhouse gas emissions. *Animal* 4: 366–376.

Yang, H. & Pollitt, M. 2010. The necessity of distinguishing weak and strong disposability among undesirable outputs in DEA: environmental performance of Chinese coal-fired power plants. *Energy Policy* 38: 4440–4444.

You, S. & Yan, H. 2011. A new approach in modelling undesirable output in DEA model. *Journal of the Operational Research Society* 62: 2146–2156.

Appendix

Statistics, input, desirable output and undesirable output data per system per year

System	Year	Number of cows	Inputs				Desirable output	Undesirable output
			Home-grown feed (t DM)	Purchased feed (t DM)	Land use (ha)	N fertilizer (t N)	Energy-corrected milk (t)	GHGE (t CO_2-eq.)
LFC	2004	43	156.9	300.9	17.9	2.8	356.9	373.8
	2005	47	176.2	391.5	24.7	2.4	414.4	421.6
	2006	49	192.2	402.7	19.6	3.1	430.7	438.4
	2007	50	211.6	427.2	30.3	3.1	456.8	461.6
	2008	50	207.2	426.6	28.6	3.3	482.3	468.2
	2009	50	188.4	425.9	26.1	3.0	473.6	541.4
	2010	51	198.9	424.4	30.1	3.2	478.2	514.8
	Mean	48	190.2	399.9	25.3	3.0	441.8	460.0
	SD	2.8	18.9	45.9	5.0	0.3	45.2	56.4
LFS	2004	37	163.3	304.6	18.8	2.9	393.2	380.0
	2005	36	159.1	346.7	22.2	2.2	401.1	390.4
	2006	40	194.6	401.2	19.9	3.1	482.7	419.6
	2007	44	216.8	424.4	31.0	3.2	494.2	442.5
	2008	44	204.8	403.0	28.1	3.2	537.7	449.0
	2009	48	193.9	425.0	26.5	3.0	549.6	533.5
	2010	53	219.6	446.1	32.8	3.4	641.4	547.6
	Mean	43	193.2	393.0	25.6	3.0	500.0	451.8
	SD	6.0	24.0	49.9	5.5	0.4	87.0	65.7
HFC	2004	43	209.4	220.6	25.9	4.3	339.9	424.0
	2005	49	256.2	247.3	37.7	3.7	384.7	478.4
	2006	52	239.3	237.8	27.7	4.3	379.3	499.3
	2007	53	275.1	259.6	41.1	5.0	381.4	515.5
	2008	54	290.6	284.1	43.4	4.6	395.2	530.1
	2009	55	278.3	287.3	41.2	4.8	454.3	605.9
	2010	55	284.6	293.5	45.7	4.3	411.1	571.4
	Mean	51	261.9	261.5	37.5	4.4	392.3	517.8
	SD	4.4	29.1	27.8	7.7	0.4	34.9	59.8
HFS	2004	42	221.4	214.9	27.3	4.5	408.6	427.0
	2005	47	253.6	226.7	37.3	3.6	409.0	451.0
	2006	47	226.6	224.0	26.3	4.1	380.6	453.6
	2007	47	255.6	230.8	38.2	4.6	385.1	466.9
	2008	51	287.5	263.7	43.0	4.6	451.9	513.5
	2009	55	279.6	268.4	41.4	4.8	477.1	585.7
	2010	55	304.0	272.2	48.3	4.5	541.2	588.0
	Mean	49	261.2	243.0	37.4	4.4	436.2	498.0
	SD	4.6	30.9	24.1	8.1	0.4	57.9	66.1

GHGE = greenhouse gas emissions; CO_2-eq. = carbon dioxide equivalents; HFC = high-forage control; HFS = high-forage select; LFC = low-forage control; LFS = low-forage select; SD = standard deviation; DM = Dry Matter; N = nitrogen; t = tonnes