Benchmarking and Determinants of Technical and Economic Performances of Dutch Dairy Farms

Thesis submitted in fulfilment of the requirements for the degree of MSc. Management, Economics and Consumer Studies

Department of Business Economics (BEC),
Benchmarking and Determinants of Technical and Economic Performances of Dutch Dairy Farms

by

Ernest Dandi
MSc Management, Economics and Consumer Studies (MME)
Thesis Business Economics
Course Code: BEC-80433

Supervised by
Professor Mariska van der Voort
Business Economics Chair group
Wageningen University

Wageningen UR- April 2017
Acknowledgements

First and foremost, I glorify God Almighty who gave us strength, knowledge, and understanding to come this far. But for His grace and blessings, we could not have accomplished the assignment.

My next special and profound appreciation is dedicated to my Supervisor, Professor Mariska van der Voort for her unflinching support, guidance, encouragement, and especially her teachable and patient attitudes that always kept me on the right track. I strongly believe I would not have been able to complete my degree without your support.

Importantly, I wish to appreciate my noble sponsors, the Netherlands Fellowship Programme (NFP) for sponsoring me throughout this quality education programme in the Netherlands. May God widen the scope of NFP, and the Netherlands!

Furthermore, I wish to express my sincere appreciation to my family, and my dear friends especially James Mutambo, Engida Tadesse, Ahoa Emmanuel and Baibing Yang. Your encouragement, support and motivation were wonderful and inspiring.

Finally, I wish to express my heartfelt gratitude to Elizabeth Konadu-Kusi for your motivation, financial and prayer support.
Table of Contents

Abstract................................................................................................................................. 1

Chapter 1 Introduction, Objective and Outline ..................................................................... 2

1. Introduction....................................................................................................................... 2

1.2 Research Outline ........................................................................................................... 5

Chapter 2 Literature Review.................................................................................................. 6

2. Benchmarking.................................................................................................................. 6

2.1 Measures of Farm Performance .................................................................................... 7

2.2 Methods of Benchmarking ........................................................................................... 8

2.2.1 The Excellence Business Models............................................................................. 9

2.2.2 Database Benchmarking ........................................................................................... 9

2.2.3 Focus Group Benchmarking .................................................................................... 10

2.2.4 One to one benchmarking ....................................................................................... 10

2.2.5 Statistical Matching Method ..................................................................................... 10

2.2.6 Artificial Neural Networks (ANN) ........................................................................... 11

2.2.7 Certification Benchmarking ..................................................................................... 11

2.2.8 Efficiency Methods ................................................................................................. 11

2.3 Optimal benchmarking methods .................................................................................. 13

2.3.1 Combining Balanced Scorecard (BSC) and DEA ................................................... 13

2.3.2 Combining DEA with KPIs ..................................................................................... 13

2.3.3 Combining Efficiency analysis with Cluster analysis .............................................. 14

2.3.4 Stratification and Proximity-Based target selection methods .................................. 14

2.4. Determinants of Dairy Performance ........................................................................... 15

2.4.1 Short-term Determinants of Benchmarking ............................................................. 16

2.4.2 Long-term Determinants of Benchmarking ............................................................ 18

Chapter 3 Materials and Methods ...................................................................................... 21

3. Data ................................................................................................................................. 21

3.1 Efficiency Analysis ....................................................................................................... 21

3.2 Data Envelopment Analysis (DEA). ............................................................................ 23

3.2.1 The Input Oriented Constant and Variable Returns to Scale Model ....................... 24

3.3 Cluster Analysis ............................................................................................................ 26

3.4 Statistical Analysis ....................................................................................................... 27

Chapter 4 Results and Discussion....................................................................................... 29

4.1 Efficiency Analysis ....................................................................................................... 29

4.2 Cluster Analysis ............................................................................................................ 31

4.2.1 Cluster 1: Low TE, low performance in most cluster Variables ............................ 36
4.2.2 Cluster 2: Intermediate TE, intermediate performances on cluster variables................. 37
4.2.3 Cluster 3: Low TE, high performances on cluster variables................................. 38
4.2.4 Cluster 4: High TE, high performances on cluster variables.............................. 39
4.2.5 Cluster 5: High TE, intermediate performance................................................. 41
Chapter 5: Conclusion........................................................................................................ 44
References......................................................................................................................... 45
Abstract
Benchmarking the performance of decision-making units of dairy farms is essential from management and policy points of view. Efficiency methods have been extensively applied to benchmark performance among farms in the last 3 decades. Since farmers are not familiar with efficiency scores and its implications compared to partial indicators, it was useful to study how efficiency analysis can be made practical, and effectively applied for farm decision support. We propose a method of benchmarking that is more robust for effective decision support due to its enhanced discriminatory performance in this study. As a result, efficiency analysis (DEA) and cluster analysis were integrated to enhance the power of predicting dairy farm performances because of the ranking efficiency that facilitates accurate prediction of economic performances. We applied input-oriented DEA model to determine the optimal benchmark for the dairy farms. DEA determines the positions of dairy farms in the input-output space and rank farms based on their position. From the DEA analysis, the average benchmark of 0.835 and 0.864 for CRS and VRS models respectively were observed and farmers were mostly inefficient on labour than all input variables. By integrating DEA and cluster analysis we defined 5 benchmark groups with similar characteristics which determine economic performance among dairy farms. We categorized performance profiles into; Group with high TE and best performances in most cluster variables; Group with high TE and intermediate performances in most cluster variables; Group with low TE and worst performances in most cluster variables; Groups with low TE and intermediate performance in some cluster variables; and Group with intermediated performances in TE and cluster variables. Findings from this study indicate that combining DEA with cluster analysis enhances the ability to accurately predict the performance among benchmark groups of dairy farms and helps to assess improvements better. The group by group regression analysis showed that variables with long-term impacts on economic performance had a higher impact in predicting TE than short-term determinants. Also, herd size had no significant impact on milk yield.

Key Words: Benchmarking, Technical efficiency(TE), Economic efficiency (EE), Determinants of Dairy Benchmarks, Cluster analysis
Chapter 1 Introduction, Objective and Outline

1. Introduction

Performance measurements are crucial for effective management and long-term survival of dairy farms (Frank and Collis, 2003). Kaplan and Norton (1992) defined performance measurement as the process of quantifying an organisation’s history, determining the organisation’s current position within business society, and creating strategies and overall vision for the future. Evaluating farm performance triggers innovation through strategic thinking and resource planning (EIP-AGRI, 2017; McAdam and Kelly, 2002). A common practice to improve economic performance among dairy farms is to benchmark the relative performance against leading farms in the industry (Stapenhurst, 2009; Ray et al., 2015). Among strategic farm management tools, benchmarking is most important and it is positively correlated with farm performance (EIP-AGRI, 2017; Smeltzer and Carr, 1999), and supports farm productivity and sustainability improvement performance (Poppe et al., 2013). An advantage of benchmarking is that it follows imitation and adapting to best-recommended practices than purely on research measure or invention (EIP-AGRI, 2017). Comparing farm performance against leading farms enables farms operating below optimal performance levels to improve performance which may result in higher farm income and business competitiveness (Franks and Collis, 2003). Studies indicate that dairy farms that attain high farm profits apply benchmarking method (Wilson, 2001). Therefore, benchmarking is useful for dairy farms to improve on quality initiatives and stimulates innovation among farms. As improvements in dairy performance are indispensable to revitalizing dairy sector and economic development (ZuivelNL, 2015), benchmarking individual farms is essential from practical and economic viewpoints (Carbrera et al., 2010).

Various methods have been applied to benchmark performances of dairy farms in the past years (Ray et al., 2015, Begotoff, 2012). In agriculture, using partial indicators to benchmark farm performance is a common practice (van Meensel, 2011; Dawkins et al., 2007, Wilson et al., 2005). Partial indicators include feed conversion (kg feed per kg milk produced), the amount of milk per cow and the well-known financial ratios (Edwards et al., 2015; Begotoff, 2012). Although using partial indicators is convenient and easy to communicate to farmers (van Meensel, 2011), they yield wrong results when analysing benchmark farms on specific indicators (Ray et al., 2015). Studies have indicated shortcomings when farms are benchmarked on key performance indicators (van Meensel, 2011; Lee, 2012; Herington and Guilding, 2008). It has been established that placing emphasis on financial indicators may
result in the failure to cover multiple dimensions impacting performance (Herington & Guilding, 2008). As a farm cannot become best performing farm in all KPI’s, may result in comparing farms against unrealistic benchmarks. What is more, partial indicators relying on the assumption that all dairy farms exhibit constant returns to scale (CRS), means inputs and outputs can be scaled up at equal rates by all farms. The assumption may not apply in all practical situations making it impossible to compare a small dairy farm with a large dairy farm. As more resources may be needed by a small farm to achieve the high level of specific partial indicator, lack of adequate capital may hinder the capacity to scale up productive performance. Besides, a large farm may be constrained by coordination and control problem. Therefore, it is important to benchmark dairy farms with a method that considers the complete use of all resources and outputs and account for factors that are beyond farmers’ control (Begotoff, 2012; EIP-AGRI, 2017). In view of the diverse conditions in the dairy industry, factors determining dairy economic performance should be assessed simultaneously in an unbiased fashion. The aim of farm management is reaching a superior long-term farm performance (Gloy et al., 2002), analysing the relationship between technical and economic factors of production on farm performance will be useful for effective extension and policy decision support. It is important to analyse how to prioritise and customise relevant dairy performance influencing factors for benchmarking and determine improvement profiles for farms.

To overcome the shortcomings of PKI’s, researchers have found efficiency analysis as the most important method of benchmarking to measure the economic performance of farms (Cabrera et al., 2010; Hansson, 2007). The appeal with efficiency method lies in the ability to assess inputs and outputs simultaneously (Tremblay, et al., 2015), and compares input-output combinations against performance benchmark (van der Voort, 2015). Besides, studies show that improvements in technical efficiency are the key factor for the survival of dairy farms in traditional production areas (Tauer 2001; Alverarez et al., 2008). As economic efficiencies vary across dairy farms indicate farms can further improve their performance (Franks and Collis, 2003; Gloy 2002). However, there is no clear-cut norms for defining the maximum output achievable from a given set of inputs (Ray et al., 2015). Improving performance requires the availability of farm-specific benchmarks which serves as guides to the use of recommended practices and targets since farm and site-specific characteristics outside the control of the farmer influence the economic, environmental and societal performance (EIP-AGRI, 2017). Dairy farmers must identify and target benchmarks that are truly representative
of their industry. Achieving uniform performance level among farms require a knowledge of best achievable performance level.

When analysing farm efficiencies, two broad conventional methods used are the non-parametric Data Envelopment Analysis (DEA) and Stochastic Frontier Analysis (SFA) (Coelli et al., 2005; Kumbhakar & Lovell, 2000; Carbrera et al., 2010; Shaneth, et al., 2009; Stokes et al., 2007). DEA is widely applied in modern benchmarking works in management and economic studies (Shaneth, et al., 2009; Stokes et al., 2007). The strength of DEA lies in its ability to handle the multi-product nature evidenced in dairy farms simultaneously (Kodabakhshi and Asgharian, 2008) and producing more robust results (Sharma et al., 1999). DEA does not require parametric specification of any functional form when constructing the production frontier (Coelli et al., 2005). There are therefore no unnecessary restrictions regarding functional form that impinge upon efficiency analysis and disrupt the efficiency results. Besides, DEA provides quantitative insight into the adjustments needed to correct inefficiencies among worse performing dairy farms (Ray et al., 2015; Coelli et al., 2005). Despite its wide application, the methodology does not take into account the effects of exogenous variables as staff motivation, animal genetics, managerial competence, etc, whose effects on dairy economic performance are not readily quantifiable. It is difficult to relate directly the impacts of stochastic factors on dairy economic performance at least in the short term as there are a lot of interactions between conversions of inputs into outputs. Nonetheless, it is difficult for inefficient farms to become efficient by benchmarking against a target with different input composition. Identifying appropriate benchmarks that are similar in input composition is handy for inefficient farms to imitate targets (Shaneth et al., 2009). In view of these drawbacks, the current paper aims to study how frontier estimation methods such as DEA can be applied for benchmarking and dairy decision support. Essentially, the study will investigate which methods can complement DEA for benchmarking in more practical way. Besides, the study aims to benchmark Dutch dairy farm performances by investigating the determinants of benchmarking impacting dairy performance in the short and long term.
1.2 Research Outline

This paper is organised into five parts. Chapter two consists of literature part. Based on the literature study, I put forward points of attention that must be considered when analyzing benchmarking and determinants of technical and economic performances of dairy farms. Findings from the literature review contributed to answering research question 1 and 2.

Chapter 3 considers the materials and methods part. It discusses DEA, the efficiency models and cluster analysis of benchmarking. The chapter highlights the dataset and dairy farm selection criteria. The variables selected for the DEA and cluster analysis will be discussed. Finally, statistical analysis that will be done on the results is elaborated.

Chapter 4 talks about results and discussions. The chapter discusses the results based on the methods in Chapter 3. We discuss the results considering findings from other similar works that have been done on the subject. Chapter 5 considers conclusion. The chapter discusses whether the research answered the research questions and the problem mentioned in the introduction. We describe the implications of this study for farmers and policy makers and possible recommendations on how farm-specific decision support on farm performances can be further developed.
Chapter 2 Literature Review

2. Benchmarking

Benchmarking is a vital part of the farm improvement strategy. This is because by benchmarking, farms escape investment sinking mistakes that other decision-making units made as well as avoid duplicating the efforts of other trailing dairy farms. Benchmarking shows how competitive the farm is, where to focus performance activities as well as the optimal performance levels on which to gauge performance (Stapenhurst, 2009). Benchmarking helps to facilitate decision making (learning and coordination) and control of activities (motivation) (Begotoft, 2012). Though not usually mentioned as an objective of carrying out a benchmarking exercise, benchmarking can be a useful tool to identify the competitive weak points of the competitor (Stapenhurst, 2009)

Different authors have offered an alternative and complementing definitions of benchmarking. EIP-AGRI (2017) defined benchmarking as improving the performance of a farm, for example, by comparing with peers, learning from others and identifying actions. The definition of EIP-AGRI’s agrees with Franks and Collis (2003) that benchmarking is not a radical approach for managers to improve performance. Benchmarking is a conscious and systematic approach of assessing farm performance.

Stapenhurst (2009) defined benchmarking as the method of measuring and improving organisational performance by comparing ourselves with the best. The definition focuses benchmarking on two (2) phases. The first phase describes farm performance. It explains quantifying the performance level of participants, identifying the gap between farms, and quantifying what farms stand to gain should all operate at the optimal benchmark. The second phase involves changing practices to improve the performance of the farm. Begotofts (2012) defined benchmarking as a managerial tool that improves performance by identifying and applying best-documented practices. Inferring from these definitions shows that farm managers analyse, measure and compare their productive performance regarding products, processes externally with best performing farms or best-in-class companies and internally with other operations within their own farms that perform similar activities. To summarise, benchmarking is improving farm performance by identifying best practice, measuring performance against best practice and establishing affiliations between best practice farms (peers) and worse performing farms in order for worse performing farms to identify and adopt best practices.
2.1 Measures of Farm Performance

Several measures give an indication of farm performances (Bojnec and Latruffe, 2008). Bojnec and Latruffe (2013) studied farm size, agricultural subsidies, and farm performance and applied profitability and efficiency as the measures of farm performance. Profitability is based on the firm’s ability to generate sufficient earnings to cover the farm’s operational costs. Regarding profitability, they found cost and revenue ratios to be most important indicators. Empirical studies have established that cost of producing milk is a major indicator of economic sustainability among dairy farms (van Chalker, 2005), and means of measuring farm competitiveness in both factor and product market (Thorne, 2004). A farm’s efficiency is based on a comparison between the observed inputs on one hand and optimal outputs on the other hand. Although efficiency is an important condition for determining effectiveness, it is not a sufficient one (Ray et al., 2015). This can be explained by considering the requirements of particular preference function. It is possible for the inefficient farm to be much better than a fully efficient farm. It is therefore not enough for dairy farms to lie on the production frontier, farms should strive to lie at the proper place on the production frontier. Wilson et al., (2005) suggest benchmarking should focus on key variable influencing productivity, profitability, liquidity, and solvency.

Effectiveness describes a firm’s ability to achieve or exceed a previously defined performance target. The focus of effectiveness is on the output of the productive process (Begotoff, 2012). When applied to evaluate performance, it focuses mainly on output but not the inputs employed in the productive process. Moreover, as effectiveness is judged on the farms objective, the selection of an appropriate objective or the preferences against which performance should be measured is arbitrary (Ray et al., 2015). This problem is dire since farms are pursuing diverse objectives at a time.

Productivity has been the most widely applied measure of performance. The productivity of a farm explains the ratio of the amount of output and input employed in the productive process. A disadvantage of productivity is the tendency to ignoring contributions of other factors, and may not reveal much information on longer-term farm economic sustainability (Wilson et al., 2005). Consequently, it is possible that poor performing firms are shown in more favourable light (Ray et al., 2015).

As many factors influence performance, selection of an appropriate measure is important. First, the analyst must consider controllability of resources, time in relation to productive
resources, the intended use of efficiency results, ease of interpretation of the results, data availability and ease of computation (Begotof, 2012)

2.2 Methods ofBenchmarking
Various methods have been applied to benchmark farm performances and inspire improvements across different fields (Ray et al., 2015; Stapenhurst, 2009). Knowledge of the different methods of benchmarking and understanding them is both important and useful. By knowing and understanding different benchmarking methods, farmers and analysts escape the danger of trying so hard to adopt methods that are technically and practically inappropriate to solve a problem of the farm. Moreover, it facilitates searching and applying the most effective and practically plausible methods to achieve the requisite performance target. Talluri (2000) suggested that for successful evaluation, benchmarking method must simultaneously analyse and integrate several key performance measures to identify best practices. Stapenhurst (2009) outlined seven broad methods of benchmarking. These are public domain benchmarking, one-to-one benchmarking, review benchmarking, database benchmarking, trial benchmarking, survey benchmarking, and business excellence benchmarking. Begotof (2012) and Ray et al (2015) categorised benchmarking methods as traditional, and modern which are essentially efficiency methods since they have been extensively applied for ranking farms in recent literature in the past three decades (Coelli et al., 2005). Efficiency methods are further categorized into parametric and non-parametric methods (Coelli et al., 2005). The traditional methods like excellence business models benchmark farms on some key performance indicators (KPI’s). Literature indicates that combining different methods for benchmarking is more effective (Rahimi and Behmanesh, 2012; Stapenhurst, 2009; Adler et al., 2002) since a single method of benchmarking cannot address farm situations, objective, and preferences always. Therefore, an optimal approach is to be able to blend multidimensional information in a useful and effective way.

Various reasons influence the choice of an appropriate method for performance measurement (Stapenhurst, 2009). When deciding on a suitable method, the benchmarker should consider first, confidentiality requirements. Confidentiality requirement implies the extent to which data can be shared openly, shared but anonymised or not shared; second whether the participant knows which organizations are best performers or whether comparative levels of performance need to be ascertained; third the scope of study; fourth whether a benchmarking club already exists with a similar scope of the proposed study; and lastly experience of the organization with benchmarking.
2.2.1 The Excellence Business Models
The excellence business models set performance criteria on all key aspects of successful farms against which worse performing farms will be benchmarked (Drury, 2013; Wongrassamee, 2003). The extent that dairy farms implement and adhere to these criteria reflects the farm’s success or failure. The excellence business models overcome the limitations of conventional benchmarking using key performance indicators such as the use of lagging metrics, not incorporated into farm’s strategy, difficulty to implement in practice as it tends to be inflexible and fragmented, contradicting accepted performance improvement and failing to recognize customer performance. Excellence business models provide a mechanism to compare the performance of any group of organisations by scoring against a standard. The results are then compared. Direct comparisons are seldom done between companies. Reasons are that there might be differences in the way different organisations will adhere to the key performance criteria and based on the corporate objectives, one may attach different degrees of importance to certain performance criteria in the model and not the other. The major advantage of business models lies in its effectiveness in reviewing the overall performance of the company and identifying loopholes in process of work. However, the European Foundation Quality Model like many other business models such as the balanced scorecard may be too ambiguous and lack theoretical underpinnings or empirical support or validations (Drury, 2015). There might not be enough evidence to prove the relation between non-financial and future financial performance. Another important criticism relates to the failure to include very important issues of universal concern, importantly like environmental impact and social responsibilities which are critical factor for success.

2.2.2 Database Benchmarking
Database benchmarking is based on sharing data at farm level between systems and moving data electronically in a seamless way (Poppe and van Asseldonk, 2015). The method involves an independent consultant who builds up database of performance levels of farms over time. As new participants join the study, their data is added to the existing data and performance level compared with the participating farms. Database benchmarking has immense potential to increase farm productivity and sustainability performance in agriculture as well as stimulating farmers’ interest in benchmarking due to many advantages provided through information sharing (EIP-AGRI, 2017). The government of UK applied this method to identify competitive practices in the use information and communication technologies in France, Sweden, Germany, the USA and Japan (DTI, 2000). The method has received many commendations among companies for its advantages in terms of cost, time and access to
partner companies. The drawbacks of this method rest with the participating farm as there is no guarantee that their data is compared to farms with similar characteristic. The authenticity of the results depends very much on the competence of the consultant and data accuracy. The cost of benchmarking can be high for farms in terms of participating in the operations and time of data collection.

2.2.3 Focus Group Benchmarking
By this method, data is sourced from groups of farms within similar production enterprises for benchmarking. The method has much in common with database benchmarking when the analyst builds a database and keeps information taken from farmers. Focus group benchmarking stimulates some impetus to encourage farmers into making a change and adopting new practices. EIP-AGRI (2017) applied this method to study benchmarking of farm productivity and sustainability performance among some EU countries. It was found that the method is very useful since benchmarks are discussed on regular bases individually between a farmer and his or her advisor or in peer groups among farmers in 80% of the EU countries.

2.2.4 One to one benchmarking
The method involves identifying which farm is best at performing an aspect of the productive process which the farmer wants to improve, visiting the farm in order to ascertain their level of performance and learn how they achieve the performance level and adopt them where necessary (Fernandez et al., 2001). The study performed by original Xerox and Kodak Rochester plant in the in the early 1990’s is an example of one to one benchmarking method. An advantage of the one to one benchmarking lies in the ability to overcome the challenges of integration such as the transferability and diversity. The cladistic approach as an example is able to contain and represent different organizations without regard to how and when they were formed and identified.

2.2.5 Statistical Matching Method
Dolman et al., (2014) applied statistical matching technique to benchmark the economic, environmental and social performance of Dutch dairy farms. Statistical matching allows forming benchmark groups with similar characteristics that affects farm sustainability (Vrolijk et al., 2005). Statistical matching is particularly suitable with quantitative variables (Andridge and Little, 2010). The methodology establishes a (conditional) independence between the variables never jointly observed given the common variables. Compared to cluster analysis, statistical matching is less powerful and may be unnecessary since the
outcome is a priori. Besides, a conditional independence is produced for the variables not jointly observed although they may be conditionally dependent (Rassler, 2012). The validity of the matching results depends on the accuracy of the assumptions of the relationship between the variables that are unique to the input files.

2.2.6 Artificial Neural Networks (ANN)
Artificial neural network method is a flexible mathematical structure which can identify complex nonlinear associations between inputs and output data set. Researchers have found this method to be effective especially in problems where the processes are difficult to describe using physical equations, and prediction is more important than explanation (Rihimi and Behmanesh, 2012, Samoilenko, et al., 2010, Vadani et al., 2012). An effective application of ANN in data analysis is the multilayer perceptron (MLP). The multilayer perceptron is a nonlinear neural network model that can be applied to approximate any function with high degree of accuracy. ANN has been applied widely in the energy sector for predicting the production and consumption of energy (Yalcintas & Aytun-Ozturk, 2007). As a classification tool, artificial neural network is found to be more suitable and stable than regression (Emrouznejad and Anouze, 2010). ANN allows farms to discover meaningful and previously hidden information from a large database.

2.2.7 Certification Benchmarking
Certification benchmarking is essential for quality assurance and environmental sustainability. In this method, the participating farms are required to follow or meet a certain environmental and quality standard. This implies the driving force is not directly the consumer, but other actors in the production chain. This method is important for increasing awareness of environmental issues and has been employed as a key benchmark method among crop production enterprises via internet (Udo de Haes & De Snoo, 1997)

2.2.8 Efficiency Methods
SFA and DEA are the frequently used efficiency methods of benchmarking (Carbrera, et al., 2010; Coelli et al., 2005; Ray et al., 2015). Proposers of this method deem SFA most appropriate benchmark method for most agricultural enterprises because of its ability to handle stochastic noise, accommodate traditional hypothesis testing and allow for single step estimate of the efficiency effects (Stokes et al., 2007; Kumbhakar and Lovell, 2000). Stochastic frontier estimation makes an explicit specification of the functional form of the production process and applies maximum likelihood estimation procedure to estimate the
parameters of the parametric model (Bojnec and Latruffe, 2013). The models are characterised by a priori specification of the functional form, with the exception of the parameters which are estimated from the dataset. The methodology reflects the importance of the different cost drivers or parameters in noise and efficiency distributions (Ray et al., 2015). Moreover, the SFA allows estimating elasticities or other features of the technologies from the model. A major drawback of the SFA relates to difficulties when estimating efficiency in a multiple-product industry (Coelli and Perelman, 1996) and requiring an a priori functional specification by the analyst which may be arbitrary (van der Voort, 2015; Jaforullah and Devin, 1996). The validity of the estimated efficiency results become tentative since it depends on the appropriateness of the functional form (Begotoft, 2012).

DEA has flexible functional form and allows ranking efficiencies of decision-making units with multiple inputs and outputs (Coelli et al., 2005; Fraser and Cordina, 1999; Khodabakshi and Asgharian, 2008). Studies have sought to improve on the differential capabilities of DEA methodology. Adler et al., (2002) outlines six distinct ranking approaches of DEA. These are cross-efficiency matrix, in which the units are self and peer evaluated; the super-efficiency method which ranks through exclusion of unit being scored from dual linear program and an analysis of the change in the Pareto frontier; benchmarking in which a unit is highly ranked when chosen as useful target for many other units; multivariate statistical techniques which are applied next to DEA dichotomic classification; ranking inefficient units through proportional measures of inefficiency; the sixth approach requires collection of additional preferential information from relevant decision-makers and combine multi-criteria decision methodologies with DEA. Stochastic DEA combines the flexible properties of DEA and allows the possibility that some of the data variations may be attributable to noise and requires only most of the points to be enveloped (Khodabkhshi et al., 2010; Cooper et al., 2004; Begotoft, 2012). The super-efficient model of stochastic DEA allows analysts to rank efficient decision-making units and accounts for possible uncertainty in inputs and outputs (Adler et al., 2002). Although each of these methods are efficient in a special way, no one approach can be prescribed as a complete solution to the question of ranking decision-making units (DMU’s). Table1 below give the classes of frontier benchmarking methods
Table: 1 Classes of efficiency benchmarking methods

<table>
<thead>
<tr>
<th></th>
<th>Deterministic</th>
<th>Stochastic</th>
</tr>
</thead>
<tbody>
<tr>
<td>Parametric</td>
<td>Corrected Ordinary Least Square (CORLS), (Ray et al., 2015)</td>
<td>Stochastic Frontier Analysis (SFA), (Kumbhakar and Lovel, 2000)</td>
</tr>
<tr>
<td>Nonparametric</td>
<td>Data Envelopment Analysis (DEA) (Coelli et al., 2005)</td>
<td>Stochastic Data Envelopment, SDEA. Begotofii(2012)</td>
</tr>
</tbody>
</table>

2.3 Optimal benchmarking methods

2.3.1 Combining Balanced Scorecard (BSC) and DEA.
Traditional performance measurements place much emphasis on financial indicators and this can potentially cause managers to under-invest in nonfinancial components that are important for achieving long-term farm success and thus fails to deal with multiple dimensions that impact performance (Lee, 2012; McPhail et al., 2008). Therefore, the BSC is very important in this instance since it provides a comprehensive performance measurement of both financial and non-financial perspectives of the farm (Cokins, 2005). A major advantage of the BSC rests in its ability to translate the farm mission into goals, aligning individual farm goals, actions, and performance measures and measuring process related to goal attainment (Frigo and Krumwiede, 2000). The downside of BSC is that it does not provide a common measurement and it lacks a standardised baseline or benchmark to compare performance (Banker et al., 2005). To overcome the limitations of BSC, researchers have found DEA to be most useful (Najafi et al., 2009; Chen and Chen, 2007). Lee, 2012 applied the BSC- DEA to study kitchen employee performance measure and found the integration of the BSC and DEA was an effective performance measurement tool for evaluating a kitchen’s efficiency, identifying the best decision-making units. The integration allowed chefs and managers to identify specific areas that needed to be improved and proposed solutions on improvements in efficiency.

2.3.2 Combining DEA with KPIs
Efficiency methods of benchmarking show farm-specific economic improvement paths and margins. They determine an explicit diagnostic on the farm-specific input-output transformation by comparing it to the frontier which helps to differentiate the various improvement paths. Although frontier analysis provides an improved benchmarks, dairy farmers are generally not familiar with efficiency scores, and communicating findings to farmers may not be straightforward (Van Meensel, 2011). Therefore, efficiency score alone may not be useful for farm decision support. In this regard, DEA adds little or no value to
extension application. By combining DEA with traditional KPIs, frontier methods can provide an added value to traditional decision support (de Resende Horta et al., 2009, van Meensel, 2011). The optimal benchmarks for technical efficiency are the targets on the frontier. The targets are the linear coordinates of peer farms that define points of efficient projections on the frontier for an inefficient farm (Coelli et al., 2005). Technical efficiency score permits comparisons of farm specific information on input-output information with a relevant benchmark. By estimating allocative efficiency, the production frontier methods permit detecting farm specific benchmarks that correspond to the most effective cost minimizing input combinations. As farms with similar efficiency scores may have different economic performance due to absolute price differences, additional benchmarks need to be assessed when identifying improvement margins for absolute prices. KPIs can be used to describe the assessed improvement paths.

2.3.3 Combining Efficiency analysis with Cluster analysis
Integrating DEA with cluster analysis has been applied for benchmarking on diverse subjects (van der Voort, 2015; Emrouznejad and Anouze, 2010; Samoilenko and Osei, 2010). Samoilenko and Osei-Bryson (2008) developed an integrated methodology consisting of DEA, cluster analysis and decision tree that allows an increased discriminatory power of DEA in the presence of heterogeneity of sample of observations. Cluster analysis permits testing for naturally occurring subsets in the sample whereas DEA allows determining both relative efficiencies of decision-making units and relative efficiencies of each subset identified in the previous step. Decision tree allows predicting the sub-specific nature of the relative efficiencies of the decision-making units. Rahimi and Behmanesh (2012) found that integrating DEA, cluster analysis, and decision tree is an effective and efficient way to improve classification and forecasting accuracy. A major shortcoming with clustering is that it cannot show the optimal benchmark level against which performance can be measured. The problem can be solved when combined with DEA analysis.

2.3.4 Stratification and Proximity- Based target selection methods
Important shortfalls of DEA regarding benchmarking can be categorized into three. First, the projected peer farm might be hypothetical and might not exist in a practical sense. Second, it is uncommon to find projected inefficient farm with multiple efficient peer farms making it difficult to benchmark multiple best practice farms simultaneously (Lim et al., 2011). Third, it is often impossible for an inefficient farm to reach the projected point in a single step. Basically, it is often rare locating a targeted farm which is most efficient and similar in input
use in a practical sense (Shaneth et al., 2009). A reasonable approach to select a realistic benchmark for inefficient farms is selecting optimal benchmarks which are similar in input use (Shaneth et al., 2009). An ultimate approach is to provide an optimal path to most inefficient farms on the frontier eventually through several times of proximity based target selection process. This process is described as the proximity based target selection method. Shaneth et al., (2009) has proposed this method in combination with reinforcement learning and self-organising map to benchmark Canadian Bank branches. These authors observed that proximity-based selection method was very practical to obtain a gradual improvement for inefficient decision-making units (banks) than previous methods. Lim et al., (2011) propose a method which follows a sequence of intermediate benchmark target advancing towards the ultimate target. The method starts by stratifying decision-making units which entail clustering them into several layers according to their relative efficiencies. Benchmarks are established across the sequence of layers. A preferable farm in the next layer is selected as the next benchmark target based on attractiveness, progress, and infeasibility.

To conclude, each method of benchmarking is effective in a peculiar way of assessing the farm. However, none of the benchmarking methods can be adjourned superior to the subject of ranking (Adler et al., 2002; Stepenhurst, 2009). Optimal and effective benchmarking can be derived by integrating one or more methods in an objective way.

2.4. Determinants of Dairy Performance
The performance of dairy farm hinges on many determinants impacting on the level of performance across the dairy sector (Tauer and Mishra, 2006; El-Osta and Johnson, 1998). Latruffe et al., (2005) grouped determinants of farm performance into farm characteristics, environmental characteristics, and socio-economic characteristics. Variables such as farm size, debt structure, and farm specialization relate to the farm characteristics. Environmental characteristics include climate, altitude and soil quality whereas the age of a farmer and type of farm relates socioeconomic characteristics. Hansson (2007) discusses determinants in terms of the level of environment influencing dairy performance as internal, external and operational. The internal environmental factors are those factors of performance that are under the direct control of the farmer at least in the short run. The external environmental factors correspond to macroeconomic factors over which farmers have no control at least in the short term. Operational environment describes the market situation over which the farm may have some control. Jafurrullah and Whiteman (1999) rather classed determinants into
controlled, which are easily quantifiable and included in DEA, and those that are beyond the control of farmers and are not easily quantifiable for evaluation in DEA.

### 2.4.1 Short-term Determinants of Benchmarking

Empirical studies have found many determinants of economic performance with short-term effects on the farm. The list of the short-term factors is not exhaustive and may differ across various enterprises in agriculture. We discuss here some of the most important short-term factors impacting the dairy economic performance.

Farm size has been found to influence the economic performances of farms (Keszthelyi, 2014; Carbera et al., 2010; Gloy et al., 2002). Findings on this determinant are mixed (Gloy et al., 2002). Conventionally, farm size is measured in terms of the value of farm assets, land size in hectares, number of cows or work units (EIP-AGRI, 2017). These measures do not reflect real economic size since they are not directly associated with farm performances. The economic size can, therefore, be expressed as the earning capacity of the farm such as standard gross margin (Keszthelyi, 2014). Carbera et al., (2010) studied factors determining technical efficiency among 273 dairy farms in Wisconsin, USA using the SFA. Farm size, expressed as number of cows, was found as the most important determinant of dairy farm performance. Bailey et al., (1997) used an economic simulation model and production plan to analyse the impact of economies of scale on dairy profitability and observed that only the 500 and 1000 unit farms, the largest range farms proved viable. Many studies have found a positive relationship between farm size and technical efficiency (Wronski et al., 2007; van Passel et al., 2006; Mishra and Morehart, 2001). Annual maintenance cost per cow decreases with larger herd size, and larger farm size has sustainable efficiency. Robinson (1962) found a negative correlation between farm size and technical efficiency and identified some reasons why technical efficiency and farm size are negatively correlated. Technical efficiency (TE) describes input amount converted to producing a given amount of output. These reasons include: (1) the replacement of gains of labour division with cost as routines cause boredom and diminish creativity, (2) the reduced speed and flexibility of decision-making, and (3) increased cost of coordination.

Farm specialization is an important determinant of dairy farm performance (Spicka and Smukta, 2014; Latruffe et al., 2005; Kopeva and Noev, 2002). Latruffe et al., (2005) studied the impact of farm specialization on technical and scale efficiency among livestock and crop farmers in Poland and found a positive correlation between farm specialization and technical efficiency. The empirical results showed that specialized livestock farms had higher technical
efficiency than specialized crop farms. Their results agree with Spicka and Smukta (2014) who found that specialized milk producers inefficient EU regions obtain much more milk yield.

The amount of milk produced per unit measure is a major determinant of economic performance of dairy farms (Thomassen et al., 2009; Gloy et al., 2002). Studies found a positive correlation between milk produced per cow and farm financial success (Short, 2000; El-Osta and Johnson, 1998). However, a negative correlation between intensification (feed/cow) and technical inefficiency was found by El-Osta and Johnson (1998). Their results agree with the results of Alvarez et al., (2008) and Kompas and Chu (2006) for dairy farms in Spain and Australia respectively. Alvarez et al., (2008) found intensive farms closer to their cost frontier than extensive farms. Carbera et al., (2010) found a positive relationship between farm efficiency and intensification, the level of contribution of family labour in farm activities, the use of mixed ration feeding system and milking frequency. Alvarez et al., (2008) observed that use of pasture, a common practice in dairy farming negatively correlated with technical efficiency. Hansson (2007) studied the impact of strategy factors on dairy farm performance in Sweden and observed that forage machinery negatively correlated with economic efficiency.

Labour requirement and type is a key determinant of dairy performance. A negative relation between physical labour use and herd size has long been observed (Wilson, 2010). Maunder (1951) studied on physical labour use from 1947 to 1949 and found that labour requirements (hours) per cow varied from 236hrs/cow/year for herds of less than 10 cows to 135hrs/cow/year for herds of 50 cows and over. Wilson (2009) also found that labour use ranged from 42.5hrs/cow/yr with a herd size of 57 cows to 26.7hrs /cow/yr with a herd size of 146 cows. Therefore, as herd size increases, fewer labour hours are required for dairy productive activities. The type of labour used for productive activity affects dairy performance. Thomassen et al., (2009) found that a higher proportion of family labour to the total labour leads to increase in TE. The desire of family members for better economic fortunes and social welfare stimulates a greater effort towards production. The results agree with the findings of Latruffe et al., (2005) that livestock farmers use a greater proportion of family labour than the corresponding crop farmers and were more technically efficient. The reason is because family labour is in full control of the resources and technology, and they are the direct beneficiary from the farm income, they are motivated to act efficiently. Hired labour may bring in additional expertise on some tasks but there may also be shirking.
Milking frequency is another key factor influencing dairy performance. Carbrera et al., (2010) found milking frequency positively correlated with TE. Dairy farms that milked their herds more than twice a day were found with higher performance scores than those that milked their herds twice (Cabrera et al., 2010; Wronski et al., 2007; Bewley et al., 2001a; Wagner et al., 2001). Erdman and Varner (1995) had similar results when they obtained 3.5kg and 4.9kg additional milk for dairy farmers that milked their farms three (3) times and four (4) times daily, respectively. They concluded that farm intensification, milking system and, labour type were the most important determinants of dairy benchmarks. Opposing results were found from earlier studies on milking frequency. Hallan and Machado (1996) argue that there is little evidence that higher levels of facilities, machinery or equipment (milking parlours and free stall housing) are related to increasing farm performance. El-Osta and Johnson (1998) investigated the use of advanced milk parlour. They concluded that technology (using milk parlour) had no significant impact on the farm income. Rather, it was observed that amount of milk produced per cow was a key determinant of dairy TE. Genetic composition, feed quality, and health (disease incidence) management explained the differences in performance.

The management capacity of farmers, as well as the management practices implemented, is an important determinant of economic performance among dairy farms. Management capacity describes the appropriate individual characteristics and skills to tackle the right problems and opportunities in the right moment and in the right way (Galanopoulos et al., 2006). Ford and Shonkwiler (1994) found that increasing dairy managerial ability offers a greater impact on dairy economic performance than increasing the number of cows. These authors observed that debt per cow negatively correlated with dairy farm financial success. The implication is that decreasing economies of herd size will increase the technical efficiency. However, Ford and Shonkwiler (1994) also found a positive correlation between herd size and financial success and dairy management implying that dairy success on large farms can be achieved only with good financial management. The results confirm the findings from Tauer and Mishra (2006) that efficient production is more important than farm size in reducing the net cost of production.

2.4.2 Long-term Determinants of Benchmarking

The price of milk is a key determinant of dairy farms (Gloy, 2002; Geert, 2011). El-Osta and Johnson (1998) studied the financial performance of dairy farms and observed that the financial performance of a dairy farm is strongly affected by the price of milk. This is an
important factor influencing farm performance that farmers have no control. Their empirical results showed that other factors such as the cost of inputs, efficient conversion of labour, feed, and capital resources into milk all influence the performance of dairy farms and are under the control of the farmer at least in the short-run.

Jaforullah and Whiteman (1999) identified environmental stochastic determinants such as soil type, managerial skills, animal genetics, and climate as key determinants of the relative efficiency of the dairy farm but are not readily quantifiable. Environmental stochastic determinants were further classed as controllable, such as managerial expertise whereas environmental factors such as differences in geology, geography, climate and other stochastic events are uncontrollable at least in the short term.

Hansson (2007) observed that for long-term productivity growth, expansion and improved performance of a dairy farm, dairy farmers should rather think more of the farm strategy. The strategy of farm defines the long-term direction or objective of the dairy farm and the approach to achieve it (Johnson et al., 2015). The farm strategy is affected by external and operational environmental factors. The influence of the external environmental factors on farm strategy are not readily quantifiable, and are not under direct control of dairy farmers in short-term (Johnson et al., 2015; Lee et al., 1999). Johnson et al., (2015) described the external environment to correspond to macroeconomic condition, over which the farmer has no control. The operational environment relates to the market situation and the farmer may have some control. Dairy farmers need to know how they correlate with farm performance and integrate into their strategic plans to improve performance. As determinants of farm performance relate to the farm strategy, they lay the foundation for the farm success.

A microsocial environment such as farmers’ personal problems, the number of co-farmers, and the presence of successors affect performance. These factors define the nearest social environment in which the farm operates. Also, the internal and micro-social environmental factors affect the dairy farm strategy and are partly controlled by the dairy farmer. These factors are not directly correlated with the actual economic performance but have long term effects on the dairy farm. Family situation, contacts between farmers are important micro-social factors. The significance of family situation and their impact on farm performance was stressed by Gasson et al., (1988). Ohlmér et al., (1994) found that the presence of family related problems explained why the farmers could not detect farm related problems in time. This can lead to a decreased farm performance (Hasson, 2007). About internal environment,
Hasson (2007) categorized inputs into two types namely; long-run decisions about resource allocation and use, and long-term decisions about fixed cost (housing, forage machinery). Variables such as the direction and focus of the dairy production represent long-term strategic decisions about resource allocated for farm use at the farm, that is difficult to change in the short-run. It was observed that the difference in climate, soil quality (important for pasture and crop enterprises), differences in farmers’ attitude and business culture were important external factors that influence performance.

The amount of farm debt employed by the dairy farm is an important function of farm performance in the long-term (Gloy et al., 2002). Whilst more profitable farms employ more debt financing of their businesses, less or unprofitable farms use debt out of necessity.

The attitude of the dairy farmer has been found to influence farm performance in the long term (Hansson, 2007). As the attitude of a farmer reflects the commitment to adopt appropriate technology and best management practices to improve performance, the performance of dairy farm has a strong bearing on the farmer’s attitude. A study conducted by Hulten and Ohlmer (2003) emphasized the relevance of farmer’s attitude to a farm’s economic performance. It was observed that farmer’s attitude judged by the reason for going into dairy production differed between geographical areas.

To conclude, the determinants of farm economic performance falls into two broad categories as either short-term or long-term. Determinants with short-term impacts are often under the direct control of the dairy farmer. Long-term determinants are not within the internal farm environment, and farmers have very little or no influence controlling them. Their impacts on farms’ performance are not directly associated with the direct economic performance of the farm. Analysing the impacts of short-term and long-term variables is an important step for effective farm management problem diagnostics and control.
Chapter 3 Materials and Methods

3. Data

Farm accountancy data used in this study was sourced from the Flynth accounting firm. Final dataset for the study was constructed based on several selection criteria. The selection criteria helped to achieve an accurate analysis of the study. Dairy farms with observations that were truly unrepresentative of any observation in the dataset and missing data were removed from the observations. Typically, as a cost function to satisfy a cost minimising solution should be nonnegative, non-decreasing in input prices and output, homogeneous of degree one, and concave in input prices (Coelli et al., 2005, p23), farms with negative feed values were not selected. Additionally, the minimum herd size of farms considered for selection was 40 cows. Jafarullah and Whiteman (1999) selected farms that had at least 30 present dairy cows to study scale efficiency of Newzealand dairy farms. Farms with a minimum of 40 dairy cows were large and characteristic of Dutch dairy farms to study dairy performance. Also, dairy farms were selected if they derived at least 50% of their gross income from milk sold. This helped to achieve dataset with comparable farms. Hansson (2007) applied this criterion to select farms when he studied strategy factors and dairy performance among Swedish farms. Final dataset consisted of 270 farms which were used for the efficiency analysis and cluster analysis.

3.1 Efficiency Analysis

Efficiency analysis follows from the production theory (Coelli et al., 2005). Efficiency analysis compares current farm performances with an optimal performance given the current technology by determining a production frontier (Bojnec and Latruffe, 2013; Farrell, 1957). Following the production frontier, farms situated on or close to the isoquant achieve the best technical performance possible with reference to the technology. Determining efficiency level based on the maximum output achieved from a set of inputs is called the output-oriented technical efficiency whereas the ability to use the least input amount of input is called input oriented-technical efficiency (Coelli et al., 2005). The efficiency score is the optimal benchmark possible with the farm. Differences in technical efficiency are determined by incorporating explanatory variables into the efficiency model (van der Voort, 2015). Two broad methods are used to determine inefficiencies among farms which are the two-step approach and inefficient effect model. The two-step method involves regressing the estimated efficiency scores on selected explanatory variables with a linear regression model or censored Tobit model (Barnes et al., 2011; Kelly et al., 2012). The main challenge is that results from
the method are biased when input variables from efficiency analysis are correlated with second stage explanatory variables. Besides, effects of the explanatory variable are underestimated (Schmidt, 2011). Wang and Schmidt, (2002) suggested the single-step based on SFA, this inefficient effect model overcomes the above shortcomings. Figure 1 shows the technical efficiency from an input perspective. The hypothetical farm employs a certain amount of input to produce an output at a point P. The technical efficiency of the firm can be represented by the distance OQ, whilst QP represents the technical inefficiency. QP defines the excesses in input use. That is the allowable amount by which all inputs can be reduced proportionally without changing outputs quantity. The TE is expressed in percentage terms as QP/OP; the amount by which all inputs need to be reduced to achieve efficient production.

![Figure 1: Depicting Input Oriented Technical Efficiency (Source: Coelli, (1996)).](image)

The $TE_i = \frac{oQ}{oP}$, and can equally be expressed as $1 - \frac{oQ}{oP}$. The TE of a farm takes values between zero (0) and one (1) with a value of 1 a farm is technically efficient (farms that produce on the curve SS’).The allocative efficiency(AE) of the farm can be estimated if the price ratios are known. From the graph, the price ratios are represented by the isocost line AA’. The AE is represented mathematically as, $AE_i = \frac{OR}{OQ}$. RQ is the potential cost reduction that will be achieved.

The technically efficient farms dominate other farms in the sector (Begotoft, 2012). Consider two dairy farms, dairy farm A, with a production plan $(X_1, Y_1)$ and dairy farm B with a
production plan \((X_2, Y_2)\) that applies the same productive technology; The dairy farm A, can be described as dominating or technically efficient than Farm B, if Farm A employs fewer amount of productive inputs to produce much larger quantity of output than farm B. The dominant farms cannot be dominated by the other farms and they represent the best practices that are input-output combinations that cannot be further improved. Given the data on input and output combinations of all farms, DEA will identify farm A as the benchmark for farm B. Farm B can reduce its inputs per unit of output by using efficient production and farm management practices of farm A. The efficiency of B is the ratio of A’s input per unit of output of farm B. The technical inefficiency of B represents the possible reduction in inputs that farm B can achieve by adopting the best productive and management practices of farm A. The technical efficiency of farm A is 100 percent. This implies that the productive performance of farm A cannot be improved further in the context of the existing technology. Therefore, farm A reflects best practice benchmark attainable for farm B and other farms producing below the efficiency frontier.

The above represent an isolated scenario where a single output is produced from an input. The dairy situation is rather complicated than this case. Dairy farms produce many different outputs from a large number of inputs. Where there are many farms, each producing multiple outputs from multiple inputs, are considered for an efficiency analysis, the benchmark of the farm will consist of many farms unless the farm is itself producing at the optimal level and lie on the production frontier. A dairy farm usually will not exemplify best practice in producing all outputs types. In effect, a best practice benchmark of a farm may include a number of farms that exhibit best practice in producing one or more outputs. More importantly, since the input-output combinations of each farm will be unique, each farm will have a unique benchmark (Jaforrrullah and Whiteman, 1999; Begotoff, 2012). DEA identifies the best-practice farms from the optimal benchmarks and estimates the relative contributions of each benchmark. This way, the worst performing dairy farms can identify their relevant peers and emulate their productive and management practices.

3.2 Data Envelopment Analysis (DEA)

DEA uses linear programming to construct best practice piece-wise nonparametric frontier to evaluate the relative efficiency of the studied farms. Efficiency scores are estimated relative to frontier (Coelli et al., 2005, Bojnec & Latruffe, 2013). Charnes et al., (1978) gave the name ‘data envelopment analysis’ after it was introduced by Farrell (1957). Charnes et al,
originally proposed an input oriented constant return to scale model. Alternative assumptions that allow variable returns to scale and weak disposability of inputs was subsequently proposed by Banker et al., (1984). For our research, we focus on constant and variable returns to scale models since these two are the most widely applied in literature. The efficiency scores were determined with DEAP version 2.1 program.

Based on literature and data available, 1 output variable and 6 inputs variables are selected for the efficiency analysis. 7 variables were used for efficiency analysis to avoid “curse of dimensionality” which influences DEA results. Curse of dimensionality occurs when many variables are included in DEA model and results in many farms becoming efficient. In effect, Chambers et al., (1998) have proposed a rule of thumb that there should always be at least 3 times as many data observations as variables in DEA model specification as possible. The output variable is total milk produced in Euros. The input variables include Labour, defined as the total labour for family and contract labour measured in Euros; Feed defined as the total cost of purchased feed measured in Euros; Crop defined as the total of crop production related costs in euros including seed, crop protection, and fertilizer. Variable costs describe the total of veterinary expenses and energy cost measured in Euros; cow defined as the total number of cows in a herd; Land size defined as the total land size used for production. Table 2 shows the descriptive statistics of the variables included in efficiency analysis.

3.2.1 The Input Oriented Constant and Variable Returns to Scale Model

The LP based input oriented constant to scale model considers milk production in euros as the output variable. One output variable is used because most the farms are specialised dairy farms. Also, the dataset for this study shows most dairy farms generate about 90% of their farm income from milk. The model includes 1 output variable and 6 input variables as described in table 2 above. The model and notation is representative of the model used by (van der Voort (2015). The variable return to scale assumption is formulated by adding convexity constraint: $\sum_{i=0}^{N} \lambda_i \geq 0$ to the CRS model. The addition of the convexity constraint leads to the formation of a convex hull of intersecting plains which envelope the data points more tightly than the CRS model. In effect, the VRS model provides technical efficiency scores which are greater than or equal to those obtained using the CRS model. The convexity constraint ensures that an inefficient dairy farm is benchmarked against farms of similar size. Therefore a projected point on the DEA frontier is a convex combination of observed dairy farms. The CRS model is not restricted by the convexity constraint. In a CRS-DEA model, it
is possible that the performance of an inefficient dairy farm is compared to farms which are larger. In order to determine $\theta$ for the $N$ number of dairy farms, the linear programming problem must be solved $N$ times, once for each dairy farm in the sample. In our model $N$ is 270. The technical efficiency measure for the $i$th dairy farm is obtained by solving the LP problem: $\min_{\theta, \lambda} \theta$

Subject to

$$\sum_{i=1}^{N} \lambda_i y_i \geq y_i,$$

$$\sum_{i=1}^{N} \lambda_i x_{ij} \leq \theta x_{ij} \quad n = 1, ..., N,$$

$$\sum_{i=1}^{N} \lambda_i \geq 0 \quad (Valid \ under \ VRS)$$

Where $\lambda_i$, $x_{ji}$, $y_i \geq 0$; $i = 1, ..., N$; $\theta$ is the scalar of technical efficiency scores for each of $N$ dairy farms and $\lambda$ is the $N \times 1$ vector of constants; $x$ and $y$ are the input and output matrices respectively. The $\theta$ estimates satisfy $0 \leq 1$, with a value of 1 indicating the technically efficient point that lies on the frontier according to Farrell’s (1957) definition.

Cost efficiency (CE) scores are obtained by solving by 2 additional linear equations. AE are calculated based on the farm-specific prices of inputs. In the following equation $w_{j,i}$ are the respective farm specific input prices

$$\min_{\lambda, x} \sum_{i=1}^{N} w_{j,i} x_{ij}^e,$$

Subject to

$$\sum_{i=1}^{N} \lambda_i y_i \geq y_i,$$

$$\sum_{i=1}^{N} \lambda_i x_{ij}^e \leq x_{ij}, \quad j = 1, ..., N$$

$$\sum_{i=1}^{N} \lambda_i \geq 0$$
with: \( w_{j,i} \) are Labour price, Feed price, Crop inputs price, fertilizer price, variable input price, Cow price, land price respectively; \( X^c_{j,i} \) are the cost-efficient labour, cost-efficient feed use, cost-efficient crop input use, cost-efficient fertilizer, cost efficient variable inputs use, cost-efficient cow number, cost-efficient land size respectively. The CE is calculated as:

\[
CE_i = \frac{\sum_{i=1}^{N} w_{j,i} x^c_{j,i}}{\sum_{i=1}^{N} w_{j,i} x_{j,i}}
\]

The Cost Allocative Efficiency, \( AE_i = \frac{CE_i}{TE_i} \). The economic or the cost efficiency is the product of the Cost allocated efficiency and Technical efficiency.

Table 2: Descriptive statistics of variables for efficiency analysis 270 dairy farms

<table>
<thead>
<tr>
<th>Variable</th>
<th>Symbol</th>
<th>Unit</th>
<th>Mean</th>
<th>Standard Deviation</th>
<th>Min</th>
<th>Max</th>
</tr>
</thead>
<tbody>
<tr>
<td>Milk income</td>
<td>Y</td>
<td>€100 kg milk</td>
<td>42.34</td>
<td>2.15</td>
<td>33.87</td>
<td>50.51</td>
</tr>
<tr>
<td>Labour</td>
<td>X1</td>
<td>€100 kg milk</td>
<td>1.28</td>
<td>1.33</td>
<td>0.09</td>
<td>6.42</td>
</tr>
<tr>
<td>Feed</td>
<td>X2</td>
<td>€100kgMilk</td>
<td>10.87</td>
<td>2.24</td>
<td>5.98</td>
<td>17.64</td>
</tr>
<tr>
<td>Crop</td>
<td>X3</td>
<td>€100kgMilk</td>
<td>1.69</td>
<td>0.49</td>
<td>0.41</td>
<td>3.16</td>
</tr>
<tr>
<td>Variable Cost</td>
<td>X5</td>
<td>€100kgMilk</td>
<td>2.44</td>
<td>0.81</td>
<td>0.66</td>
<td>7.51</td>
</tr>
<tr>
<td>Cow</td>
<td>X6</td>
<td>Number</td>
<td>117</td>
<td>56</td>
<td>44</td>
<td>426</td>
</tr>
<tr>
<td>Land Size</td>
<td>X7</td>
<td>Hectare</td>
<td>51.71</td>
<td>23.44</td>
<td>12.89</td>
<td>154.68</td>
</tr>
</tbody>
</table>

3.3 Cluster Analysis

In this study, cluster analysis is applied to identify peer groups of farms with similar production systems. Cluster analysis contributes further to a farm-specific examination of the combination of technical and economic parameters that lead to different profiles of overall performance independent of performance in a single variable (Brotzman et al., 2015). Farms are clustered based on TE scores. Combining cluster analysis with DEA is to identify distinct improvement paths for sound farm-specific decision support. The integration allows farm specific approach of analysing the relation between input variables and farm performance. Cluster analysis makes groups based on a multivariable approach (Borcard et al., 2011). In this study, 15 dairy performance variables are used to cluster farms for benchmarking. Descriptive statistics of the 15 dairy input variables are presented in table 3. The variables are classed into short term and long term variables. R statistical package was used to compile and edit data and generate descriptive statistics of the selected variables. The optimal number of
clusters was chosen based on hierarchical Ward’s minimum variance method (Kobrich et al., 2003). Once the number of clusters is determined, final clusters were formed using the non-hierarchical K-means clustering method. K-Means is effective to find and describe patterns in data and build an explicit representation of knowledge (Rahimi and Behmanesh, 2012).

An important goal of the study was to determine the relevance of short-term and long-term determinants in explaining the variations on dairy economic performance. Multiple-linear regression analysis would be conducted on all performance variables (table 3) for each cluster.

Table 3: Descriptive statistics of Cluster Variables

<table>
<thead>
<tr>
<th>Variable</th>
<th>Explanation</th>
<th>Mean</th>
<th>SD</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Short-term Variables</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Kg FPCM_Cow</td>
<td>Kg FPC milk produced per dairy cow</td>
<td>8974</td>
<td>918</td>
</tr>
<tr>
<td>Concentrate</td>
<td>Kg concentrate consumed per 100kg of milk</td>
<td>7.64</td>
<td>1.20</td>
</tr>
<tr>
<td>Pasture size</td>
<td>Size of land planted with pasture in hectares</td>
<td>42.35</td>
<td>19.44</td>
</tr>
<tr>
<td>Cows</td>
<td>Average number of dairy cows present</td>
<td>117</td>
<td>56</td>
</tr>
<tr>
<td>Vet Cost</td>
<td>Cost of veterinary/ animal health per 100kg Milk</td>
<td>0.95</td>
<td>0.41</td>
</tr>
<tr>
<td>CULL</td>
<td>Proportion of dairy cows culled</td>
<td>30.</td>
<td>8.24</td>
</tr>
<tr>
<td>AFC</td>
<td>Age at first calving, in days</td>
<td>770</td>
<td>43.10</td>
</tr>
<tr>
<td><strong>Long-term Variables</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>B&amp;C Ctrl Cost</td>
<td>Cost of breeding and control per 100kg milk</td>
<td>0.34</td>
<td>0.13</td>
</tr>
<tr>
<td>Dep_M&amp;B</td>
<td>Cost of depreciation (Machinery &amp; Building) per 100kg of milk</td>
<td>5.41</td>
<td>3.45</td>
</tr>
<tr>
<td>Housing C.</td>
<td>Cost of bedding per 100 kg of milk produced</td>
<td>0.56</td>
<td>0.32</td>
</tr>
<tr>
<td>EnvCost</td>
<td>Environmental cost per 100kg of milk</td>
<td>0.87</td>
<td>0.99</td>
</tr>
<tr>
<td>COFD</td>
<td>Cost of Farm Debt, interest per 100kg milk</td>
<td>4.35</td>
<td>2.12</td>
</tr>
<tr>
<td>CM Price</td>
<td>Critical Milk price per 100 kg of Milk received</td>
<td>38.80</td>
<td>6.49</td>
</tr>
<tr>
<td>RealEst</td>
<td>Real estate expense per 100kg of milk</td>
<td>3.08</td>
<td>1.97</td>
</tr>
<tr>
<td>TE</td>
<td>Technical efficiency, VRS model</td>
<td>0.86</td>
<td>0.11</td>
</tr>
<tr>
<td>TE</td>
<td>Technical efficiency, CRS model</td>
<td>0.84</td>
<td>0.12</td>
</tr>
</tbody>
</table>

3.4 Statistical Analysis

The descriptive statistics for each cluster was estimated in order to characterise and compare identified clusters. The technical efficiencies project the difference in position on the input-output space and form part of the economic analysis when describing the difference in clusters. Data was first analysed with a Shapiro-Wilk test to determine the normality. Variables that are normally distributed will be tested with a one-way ANOVA to assess differences among clusters. For the non-normally distributed data, Kruskal-Wallis test was used to differentiate between technical and economic Variables among identified groups of
farms. The level of significance is set at $p < 0.05$. A Bonferroni test is applied for mean comparisons of variables of equal variances and post hoc Dunn test is applied for mean comparisons of variables without equal comparisons. ANOVA analysis allows explaining differences by comparing means between all clusters.
Chapter 4 Results and Discussion

4.1 Efficiency Analysis

The mean TE are 0.835 and 0.864 for CRS and VRS models, respectively. This means dairy farms can on average reduce their inputs use by 16.5 % (CRS) and 13.6 % (CRS) and still produce at the current level of output. TE ranges from 0.52 (inefficient) to 1 (efficient) and 0.56 to 1 for CRS and VRS models respectively. The efficiency scores indicate farms economic performance level. The DEA analysis showed an average benchmark is 0.835 and 0.864 for CRS and VRS formulations, respectively. 41(15.2%) and 61(22.6%) farms are operating at optimal TE under CRS and VRS, respectively. The distribution of CRSTE and VRS TE scores are shown in table 4 and figure 2 below.

Table 4: Summary results of DEA for the Dutch dairy farms

<table>
<thead>
<tr>
<th></th>
<th>CRS</th>
<th>VRS</th>
<th>Scale Efficiency</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Mean</strong></td>
<td>0.835</td>
<td>0.864</td>
<td>0.965</td>
</tr>
<tr>
<td><strong>Standard Dev.</strong></td>
<td>0.12</td>
<td>0.11</td>
<td>0.034</td>
</tr>
<tr>
<td><strong>Minimum</strong></td>
<td>0.52</td>
<td>0.56</td>
<td>0.78</td>
</tr>
<tr>
<td><strong>Maximum</strong></td>
<td>1</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td><strong>Number of Efficient farms</strong></td>
<td>41(15.2%)</td>
<td>61(22.6%)</td>
<td>46(17%)</td>
</tr>
<tr>
<td><strong>Proportion of Dairy Farms</strong></td>
<td>47(17.4%)</td>
<td>183(67.8%)</td>
<td>40(14.8%)</td>
</tr>
</tbody>
</table>

Figure 2 showing the distribution of Technical Efficiency of Dutch dairy farms
A mean benchmark level of 0.84 or 0.86 from the CRS and VRS formulations are not the economically optimal level achievable from the actual amount of input (economically optimum is 1). However, because 0.84 and 0.86 (CRS and VRS respectively) is feasible given the current technology and dairy husbandry practices, the maximum cannot be any lower than the feasible level achieved.

Scale efficiency results are summarized in table 4. The results are useful for two reasons. First, it gives an indication whether a farm is operating at the most productive scale or not. Besides, scale efficiency indicates whether the farm is employing the appropriate factor mix (Jiang and Sharp, 2014; Fraser and Cordina, 1999). The results show that average scale efficiency is 0.965. Hence, dairy farms are 3.5% scale inefficient. This implies many farms are operating at or near full-scale efficiency. 17.4% (47 farms) are operating at their optimal scale, 67.8 % (183 farms) are operating at sub-optimal scale and 14.8 % (40 farms) are operating at their supra-optimal scale. This implies more dairy farms should be larger than the current size given the current technology, and much improvements in TE can be achieved by eliminating the sub-optimal scale. From the study, medium sized farms (50-150 cows) exhibited decreasing returns to scale. This may be attributed to inadequate capital facilities constraining expansion. It might also be due inappropriate factor mix which might be because decision making might be quite complex for the farmers when selecting the least cost combinations. The analysis showed that farmers were mostly inefficient on labour. Possible improvements in TE among the dairy farmers can be achieved by non-radially reducing the labour costs, feed costs, crop related costs, variable costs, cows and land size by 29.9, 1.04, 3.02, 3.77, 11.14 and 9.9% respectively and still produce the same amount of milk.

Mean technical efficiency of 0.84 and 0.86 for CRS and VRS models agrees with the findings reported in similar studies. Farmers were mostly inefficient with labour (29.91%) than all input variables. It could be inferred that maximum gains in farm performance level could be made by improving on labour. From the dataset, much more insightful conclusions could be made by having labour recorded in specific units such as the number of hours per day worked on a farm. As observed from the results, the efficiency scores of the CRS are either less or equal to those from the variable returns to scale specification. The reason is, under VRS, the convexity constraint added to the model ensures inefficient dairy farms are compared to
efficient farms of similar size. To obtain a precise number of efficient farms, the number of variables should be reduced. However, considering the dataset and nature of dairy farming, excluding any of the variables included in the model will undermine the husbandry reliability of the analysis. The technical efficiency scores were estimated with aggregated costs of the input variables. Cost efficiencies are Farrell’s TE in aggregated models that use costs as inputs (Begotoft, 2012 p34-35). Therefore, by saving 16.5 % (CRS) and 13.6 % (VRS) on all input use, farmers also save 16.5% and 13.6% of costs on inputs under CRS and VRS formulations respectively. Michalickova et al., (2013) studied technical efficiency and its determinants with DEA among 83 Czech Republic dairy farms. Those authors reported a technical efficiency of 0.96 and found dairy farmers were inefficient with repairs and services (15.8%) than all input variables. van der Voort(2015) studied the relation between gastrointestinal nematodes and technical efficiency among Belgian dairy farmers with SFA and reported TE of 0.81 for FADN farms and 0.88 for TFAS farms. Jiang and Sharp (2014) studied CE of dairy farms in New Zealand and observed that CE was 83% for North Island dairy farms and 80% for South Island farms. Alvarez et al., (2008) observed a cost efficiency of 72% among extensive Spanish dairy farms and 81% among intensive dairy farms. Although farm and country-specific conditions vary and influence farm performance, an important implication is that as higher TE is possible from other regions, indicates much could be achieved than the current optimal minimum.

4.2 Cluster Analysis

The clustering was done using a normalized dataset of the 15 clustering variables. The final number of clusters were determined based on the Wards complete hierarchical clustering and principal component analysis in R software. Pearson correlation matrix of the 15 cluster variables is shown in table 6. The clustering yielded five groups which best described the data set, with the mean cophenetic distances (a measure of similarities) between dairy farms within clusters being least. The five groups were not uniformly sized and consist of 56, 83, 46, 43, 42 dairy farms which represented five distinct profiles in the input-output space. Group 2 consisted of the largest number of farms (n = 83), whilst group 5 has smallest farm size (n = 42). Detailed analysis of each five clusters can be found in table 5. A biplot describing the principal components 1 and 2 and correlations among variables are shown in figure 2. The Shapiro-Wilk test for normality showed that only the variable, KgFPCM followed a normal distribution. A Kruskal-Wallis test was performed to analyse the differences between technical and economic variables per cluster. Many of the variables
showed segregation as shown from the group means. Groups did not record a singular pattern of performance. The five groups followed five performance paths categorized into; High TE and intermediate performance (Group 5) on cluster variables, high TE and high overall performance (Group 4) on cluster variables, low TE and high performance on cluster variables (Group 3), intermediate TE and intermediate performance (Group 2), and low TE and low performance (Group 1). Detailed regression analysis on the clusters discussing the relevance of the short-term and long-term determinants in predicting TE among dairy farms is shown in table 7.
<table>
<thead>
<tr>
<th>Variable*</th>
<th>Group 1 n=56</th>
<th>Group 2 n=83</th>
<th>Group 3 n=46</th>
<th>Group 4 n=43</th>
<th>Group 5 n=42</th>
</tr>
</thead>
<tbody>
<tr>
<td>KgFCM**</td>
<td>Mean 9213 SD 966</td>
<td>Mean 9154 SD 726</td>
<td>Mean 9069 SD 832</td>
<td>Mean 9214 SD 577</td>
<td>Mean 8007 SD 1024</td>
</tr>
<tr>
<td>Kg Conc**</td>
<td>7.97a 1.50</td>
<td>7.67ab 1.10</td>
<td>7.38b 1.14</td>
<td>7.39b 0.93</td>
<td>7.65c 1.20</td>
</tr>
<tr>
<td>Pasture Size**</td>
<td>48.42a 18.65</td>
<td>36.57b 14.64</td>
<td>54.72c 21.30</td>
<td>34.11b 14.74</td>
<td>51.50d 70.21</td>
</tr>
<tr>
<td>Cow**</td>
<td>131a 60</td>
<td>102b 45</td>
<td>156c 55</td>
<td>97b 47</td>
<td>125a 150</td>
</tr>
<tr>
<td>Vet. Cost**</td>
<td>1.01a 0.42</td>
<td>0.97b 0.45</td>
<td>0.94b 0.41</td>
<td>0.90c 0.32</td>
<td>0.90c 0.48</td>
</tr>
<tr>
<td>B&amp;CtrlCost**</td>
<td>0.30a 0.11</td>
<td>0.35b 0.12</td>
<td>0.28c 0.11</td>
<td>0.35b 0.17</td>
<td>0.39d 0.12</td>
</tr>
<tr>
<td>CULL**</td>
<td>35.14a 7.13</td>
<td>29.81b 8.37</td>
<td>24.74c 5.46</td>
<td>27.79d 6.65</td>
<td>32.52e 9.06</td>
</tr>
<tr>
<td>HousnCost**</td>
<td>0.71a 0.83</td>
<td>0.96b 0.88</td>
<td>0.70a 0.92</td>
<td>1.14c 1.43</td>
<td>0.71a 0.73</td>
</tr>
<tr>
<td>ManureCost**</td>
<td>0.47a 0.41</td>
<td>0.58b 0.29</td>
<td>0.65c 0.27</td>
<td>0.49a 0.25</td>
<td>0.66c 0.40</td>
</tr>
<tr>
<td>AFC**</td>
<td>766a 41</td>
<td>763a 41</td>
<td>761a 41</td>
<td>777b 41</td>
<td>793c 56</td>
</tr>
<tr>
<td>Dep.M&amp;B**</td>
<td>4.96a 2.42</td>
<td>4.29b 2.20</td>
<td>5.68c 1.93</td>
<td>8.00d 5.01</td>
<td>4.55b 2.37</td>
</tr>
<tr>
<td>COFD**</td>
<td>5.46a 1.68</td>
<td>2.91b 1.57</td>
<td>4.96c 1.97</td>
<td>4.12d 1.67</td>
<td>4.99c 2.43</td>
</tr>
<tr>
<td>CMPrice**</td>
<td>39.17a 5.30</td>
<td>38.21a 5.47</td>
<td>36.71b 4.62</td>
<td>37.44c 7.60</td>
<td>42.01d 6.95</td>
</tr>
<tr>
<td>RealEst**</td>
<td>3.33a 2.35</td>
<td>3.58b 2.00</td>
<td>2.02c 0.83</td>
<td>3.39b 2.09</td>
<td>2.57d 1.56</td>
</tr>
<tr>
<td>TE_vrs**</td>
<td>0.77a 0.08</td>
<td>0.87b 0.08</td>
<td>0.79c 0.07</td>
<td>0.98d 0.03</td>
<td>0.94c 0.08</td>
</tr>
</tbody>
</table>

*KgFCM=Kilogram fat protein content milk produced per year; KgConc. = Kg of concentrate fed per 100kg of milk produced; Past Size=size of pasture for grazing; Cow=Number of cows in herd; VetCost=Veterinary cost; B&CtrlCost = Breeding and control cost; CULL= proportion of cows culled from the herd; Housn= cost of bedding per 100kg of milk produced; Manure cos= cost of manure per 100kg milk produced; AFC=age at first calving; Dep.M&B = cost of machinery and buildings depreciation per 100kg of milk produced; COFD= cost of farm debt representing interest payments per 100kg of milk produced; CMPrice=critical milk price per 100kg of milk produced; RealEst=real estate or building costs per 100 kg of milk; TE_vrs= TE from VRS model. * Different superscripts indicate cluster means within rows differ (P<0.05). ** Variable follows a normal distribution. *** Variable is non-normally distributed.
Table 6: Pearson Correlation matrix

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>KgFPCM</td>
<td>1</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>KgConc.</td>
<td>-0.087</td>
<td>1</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Past. Size</td>
<td>-0.139</td>
<td>-0.034</td>
<td>1</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Cows</td>
<td>-0.072</td>
<td>-0.025</td>
<td>0.776</td>
<td>1</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>VetCost</td>
<td>0.059</td>
<td>0.058</td>
<td>-0.114</td>
<td>0.133</td>
<td>1</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>B&amp;Ctrl</td>
<td>-0.165</td>
<td>0.042</td>
<td>-0.273</td>
<td>0.408</td>
<td>0.303</td>
<td>1</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>CULL</td>
<td>0.172</td>
<td>0.174</td>
<td>-0.101</td>
<td>0.058</td>
<td>0.011</td>
<td>0.015</td>
<td>1</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>HousnC</td>
<td>0.052</td>
<td>0.011</td>
<td>0.040</td>
<td>0.054</td>
<td>0.089</td>
<td>0.034</td>
<td>0.001</td>
<td>1</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Man.Costs</td>
<td>0.053</td>
<td>0.063</td>
<td>-0.188</td>
<td>0.047</td>
<td>-0.032</td>
<td>0.042</td>
<td>0.074</td>
<td>0.010</td>
<td>1</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>AFC days</td>
<td>-0.195</td>
<td>0.123</td>
<td>0.040</td>
<td>0.026</td>
<td>-0.063</td>
<td>0.104</td>
<td>0.040</td>
<td>-0.029</td>
<td>-0.005</td>
<td>1</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>COFD</td>
<td>-0.157</td>
<td>0.010</td>
<td>0.021</td>
<td>0.007</td>
<td>0.020</td>
<td>0.016</td>
<td>0.050</td>
<td>-0.039</td>
<td>0.182</td>
<td>0.043</td>
<td>1</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>CMPrice</td>
<td>-0.177</td>
<td>0.144</td>
<td>-0.168</td>
<td>0.190</td>
<td>0.044</td>
<td>0.221</td>
<td>0.099</td>
<td>0.121</td>
<td>-0.015</td>
<td>0.092</td>
<td>0.346</td>
<td>1</td>
<td></td>
<td></td>
</tr>
<tr>
<td>RealEst</td>
<td>-0.049</td>
<td>0.062</td>
<td>-0.047</td>
<td>0.148</td>
<td>-0.022</td>
<td>0.095</td>
<td>0.028</td>
<td>-0.116</td>
<td>0.074</td>
<td>0.008</td>
<td>0.039</td>
<td>0.162</td>
<td>1</td>
<td></td>
</tr>
<tr>
<td>Dep.M&amp;B</td>
<td>0.025</td>
<td>0.070</td>
<td>0.046</td>
<td>0.070</td>
<td>-0.086</td>
<td>-0.037</td>
<td>0.052</td>
<td>-0.135</td>
<td>0.283</td>
<td>0.064</td>
<td>0.209</td>
<td>-0.021</td>
<td>0.012</td>
<td>1</td>
</tr>
<tr>
<td>TE_vrs</td>
<td>-0.034</td>
<td>-0.035</td>
<td>-0.428</td>
<td>0.442</td>
<td>-0.188</td>
<td>0.171</td>
<td>0.038</td>
<td>-0.057</td>
<td>0.090</td>
<td>0.052</td>
<td>-0.113</td>
<td>0.029</td>
<td>0.110</td>
<td>0.039</td>
</tr>
</tbody>
</table>

¹Variables: KgFPC= Kg of fat protein content milk produced; KgConc. =Kg of Concentrate consumed per 100kg milk produced; Past Size = pasture size(ha) as proportion of total land size; Cow = Average number of Dairy Cows on farm; VetCosts= Veterinary cost/animal health cost per 100kg milk produced; B&Ctrl Costs= Breeding and Control per 100kg of milk produced; CULL= proportion of dairy cows culled from the herd; HousnC = euro amount spent on bedding per 100kgmilk; Manure Costs= euro amount spent per 100kg of milk produced; AFC_days =Age at first calving, expressed in days; COFD= Cost of Farm Debt, the amount of interest expense per 100kg of milk produced; CMprice = critical milk price received per 100kg of milk sold; RealEst=Cost of real estate per 100kg of milk produced; DepM&B = annual cost of depreciation per 100kgmilk; TE_vrs = TE from VRS model.
Figure 2 showing Biplot of the principal components 1 and 2 plane. The plot depicts directionality of variables in 2-dimensional view and the amount of variation explained by each variable. (Variables description in table 5)
4.2.1 Group 1: Low TE, low performance in most cluster Variables.

Group 1 had the lowest TE. This implies farms within Group 1 employ more than optimal input amounts to produce at the current level of output. The amount of concentrate feed per 100kg of milk was highest in this group. The cost of animal health (VetCost, table 5), and the cost of farm debt (COFD, table 5) which denotes charges on borrowed funds were highest. KgFCPM which represent the amount of milk produced was high (statistically like that in Group 4, Table 5). Pasture size was intermediate and cost of animal bedding was lowest. Age at first calving (AFC), breeding and control cost (B&CtrlCost), as well as the cost of housing (RealEst, table 5) were intermediate. The culling rate (CULL, table 5) of cows was highest and greater than the mean of the population. The low TE of this group can be explained by the cost of farm debt (COFD, table 5) which is highest among the groups. COFD is an essential factor of farm financial profitability. The relationship between COFD and TE have widely been debated. An already efficient farm has easy access to credit whereas as bad performing farm may not have equal access (Gloy, 2002). Again, the cost of capital will be relatively lower for high-profit farms whereas farms that experienced periods of low profit would have a high demand for debt financing. Generally, where significant correlation results between TE and COFD are observed, it tends to be negatively correlated with the performance of the farm. Davido and Latruffe (2003) found a negative correlation between debt to asset ratio and TE among individual livestock and corporate crop farmers. The low TE of Group 1 is also due to the high costs on animal health (VetCosts, table 5). Pearson’s correlation matrix (table 6) and regression results (table 7) accentuate this outcome. However, higher spending on VetCost has been found to have a positive impact on factor productivity (milk yield). As was observed, high VetCost is negatively correlated with age at first calving (AFC, table 6). The results of the correlation matrix agree with the results of the cluster analysis (statistically same as of cluster 2 and 3). The low TE of Group 1 can be associated with the high mean culling rate (CULL). A high culling rate adversely affects farm finances due to increased cost of herd replacement and breeding cost. When estimating net farm income, the expenses associated with herd replacements is included in operating expenses of a farm. Therefore, operating expenses increase with increasing cost of herd replacements. An important characteristic of the Group 1 is also the mean cost of concentrate which is highest across clusters. Feed cost occupies an integral portion of the cost of dairy production (Jiang and Sharp, 2014, Michalickova et al., 2013), and accounts for 30-35% of the overall cost structure of dairy production (Michalickova et al., 2013). The cost of concentrate (KgConc,
table 5) depicts farm intensification of the dairy. The results from the Pearson’s correlation (table 6) shows TE and cost of concentrate are negatively correlated. The result agrees with Jiang and Sharp (2014) who observed a negative correlation between feed costs and TE among dairy farms in the Czech Republic. Krpalkova et al., (2014) also observed that herds kept under longer intensive rearing periods showed lower conception rates and overall services and this can cause lower TE.

The regression analysis showed a correlation of determination of 0.233 (23.30%) and 0.30(30%) for long-term determinants (Veterinary cost, breeding and control, housing, building cost, manure cost, depreciation cost, cost of farm debt and milk price), and short-term determinants (Kg FPCM, concentrate, pasture size, cow, culling rate and age at first calving) respectively. VetCost (animal health cost) was significant and negatively correlated with TE which also confirms the results of the cluster analysis and Pearson’s correlation matrix. The outcome indicates that variables under the direct operational control of the dairy farmers (or short-term variables) accounted for greater variation in TE among dairy farms in Group 1. The relatively low R-square values realized from the regression analysis might be due to other variables of dairy economic performance that are not included in the model.

Group 1 can further improve farm performance by adopting feeding system that optimize dairy nutrition at minimum cost as possible. Given that concentrate cost consist of over 30% of dairy cost structure, farmers should aim at improving concentrate efficiency. Another area for improvement is to reduce culling rate of herd whilst addressing management issues within the farm. This can result in lowering cost of breeding and veterinary costs without adversely affecting milk yield. Farms within Group 1 should adopt financial management methods that can reduce the cost of farm debt. Based on the cluster results, an impact point is to target improved milk yield as higher milk yield will reduce the COFD/100kgmilk.

4.2.2 Group 2: Intermediate TE, intermediate performances on cluster variables

The TE of this group was higher than the mean of the general population and follows groups 4 and 5. The intermediate performance relates to kgFPCM, the cost of concentrate feed, pasture size, veterinary costs, breeding and control cost, culling rate, manure cost, housing cost and critical milk price. The mean cost of asset depreciation and cost of farm debt were lowest in the group. Herd size was intermediate but was lower than population mean. An important characteristic of Group2 relates to the cost of buildings (RealEst, table 5) which is highest across the clusters. Besides, Group 2 had the least machinery and building
depreciation (Dep_M&B, table 5). The low cost of asset depreciation can be attributed to farm assets which may be relatively old. The coefficient of determination (R-squared value) from the regression analysis for short-term and long-term determinants were 0.069 (6.9%) and 0.135 (13.5%) respectively. The cost of farm buildings (RealEst, table 5) showed significantly negative correlation with TE in this group (table 7).

An advice for farms in Group 2 is to refurbish buildings and install new equipment to enhance efficient farm operations. Dairy farms within Group 2 can also improve performance by selecting good dairy breeds. The relatively high cost of manure can be reduced by adopting internal nutrient recycling or energy generation from the manure (Dolman et al., 2014). Based on the regression, much gains in dairy economic performance can be achieved by improving farm management practices that improve farm performance regarding the long-term variables.

4.2.3 Group 3: Low TE, high performances on cluster variables.

The mean TE of Group 3 is below the mean of the entire population (Table 3), and follow Group 1. Group 3 had the best performances on the cost of concentrate feed, breeding and control cost, percentage culling, manure cost, age to first calving (AFC, table 5), and building costs which are lowest across all clusters. Although the AFC was the least, the group’s mean was far higher than optimum values identified from other studies. Studies have found an optimal mean AFC to be in the range of \( \leq 24 \) months or 720 days (Stevenson et al., 2008, Mourits et al., 1999). The group’s performance on the cost of asset depreciation, KgFPCM per cow per year, and veterinary cost were intermediate with Group 3. A major characteristic of Group 3 relates to the large scale of production, measured on herd size and pasture size which are highest across clusters. In view of the low TE of Group 3, a possible cause to the low performance of farms within the group might be attributed to coordination and control and management problems which are important causes of diseconomies of size (Edwards et al., 2015). Moreover, the low TE can be explained by the mean critical milk price (CMPrice, table 5) which is lowest across all groups and the mean of the entire population (Table 3), (Gloy et al., 2003). The results in table 5 show Group 3 spends the least cost on concentrate per 100kg milk but the highest cost on pasture size. Two implications can be drawn. First, the low cost of concentrate might be due to the substitution of concentrate feed with more of pasture. Second, considering the large scale of production, the low cost of concentrate can also be attributed to cost economies due to the large size of production. The combined effect
of short-term and long-term determinants was 0.19 (19%) and 0.20(20%) respectively in predicting TE among dairy farm. The cost of animal health (VetCost, table 5) was significant and negatively correlated with TE in this group. Herd size (Cow, table 5) was significant at 10% but with no correlation effect (table 7) on TE. The results indicate herd size had no significant impact in explaining TE among dairy farms.

As empirical studies have found pasture is negatively correlated with TE due to effects on feed inefficiency and low milk yield, improving efficiency among farms could be achieved by ensuring a good balance between the pasture and concentrate. Importantly, the group should focus on improving milk production with improved breeds of cows and optimal amount resources such as feed, and cost of operation. Selecting improved breeding herds is important for improving milk yield.

4.2.4 Group 4: High TE, high performances on cluster variables.

The mean TE is highest among groups and higher than the mean for the entire population. The mean cost of concentrate, the size of pasture, herd size, and VetCost were the least (Table 5). The mean bedding cost, breeding and control cost, culling rate, and cost of farm debt were intermediate with Group 4. The mean cost of manure, asset depreciation and AFC and KgFPCM were highest with this group. The price of milk is low and follow Group 3.

The high TE in Group 4 is due to the KgFPCM (9214kg per cow per year) which is highest across groups and greater than the mean of the entire population (table 3). This observation agrees with empirical results from other studies. Thomassen et al., (2009) found a positive correlation between amounts of milk produced to be positively correlated with technical efficiency. Group 4 had the least mean cow number among the clusters. Comparing the cluster means on KgFCPM and Cow (herd size), the empirical results shows herd size has less influence on the amount of milk produced and TE. This agrees with the regression results on Group 3 found in cluster 3(table 7). Hemme et al., (2014) studied benchmarking the cost of milk production among 46 countries and observed that milk yield and prices correlate highly with cost but not herd size. The high TE in cluster 4 can also be explained by the low pasture size which is least among the group by group means. The results from the Pearson’s correlation matrix in this study shows a negative correlation between pasture and TE. Studies on similar subjects also found a negative relationship between pasture and TE (Bargo et al., 2002, Carbrera et al., 2010, Hanson 2007) due to negative effects on feed efficiency. Hansson (2007) studied strategy factors as drivers of dairy economic performance among Swedish
dairy farms and observed that pasture machinery negatively correlated with the economic efficiency of dairy farms. The high TE of Group 4 is also due to lower cost of concentrate feed (KgConc) and veterinary or animal health costs (table, 6).

The Pearson’s correlation matrix (Table 6) showed a negative correlation between VetCost and AFC. Groups 4 had higher mean age to first calving and lowest VetCost confirming the relation observed in table 6. AFC was observed to be negatively correlated with TE following the regression analysis on Group 5. The result sides with the findings of Krpalkova et al., (2014) and Etemas and Santos (2004). Spicka and Smutka (2014) found that efficient dairy farms in EU regions spend slightly more specific livestock cost (feed, bedding, and veterinary costs) per hectare than inefficient regions. These costs generate a higher livestock factor productivity. Low AFC leads to decreased feed costs, greater cumulative production, and shorter gestation interval (average age of cow at the birth of offspring) and lower overhead costs. However, early AFC has been found to cause decreased conception rates, increased dystocia, reduced milk production per lactation, diminished longevity, and increased costs of increased nutrient density in the ration (Hoffman et al., 1999). Another important characteristic of Group 4 was the low mean culling rate (CULL, table 5) among farms. Brotzman et al., (2015) observed a negative correlation between culling rate and herd performance since replacement efforts are likely to be associated with health problems in early lactations. Lower average culling rates were found to be related to increasing herd size, higher age among herd, and higher lifetime milk production and higher productive life whereas high culling rates leads to herds decreasing in size, longer calving intervals, higher 305-d fat and protein production of cows present in the herd and higher average SCC (Nor et al., 2013). From the regression analysis, the combined effect of short-term and long-term determinants was 0.054 (5.4%) and 0.432 (43.2%) respectively in predicting TE among dairy farm. The cost of animal health (VetCost, table 5) was significant and positively correlated (0.031) with TE. Depreciation of assets (Dep.M&B, table 5) showed a significantly negative correlation (-0.003, table 7) with TE.

From table 5, economic performance can be achieved by producing at optimal VetCosts (to improve health) reducing depreciation and building costs. Studies have found a negative correlation between capital costs and dairy TE. Spicka and Smukta (2014) observed that the productivity of energy inputs and capital costs were the key determinants of specialized dairy farm efficiency. Capital costs consist of depreciation, rent paid, interest, machinery and
buildings costs, taxes and other charges on land and buildings. Managing manure constitutes an essential cost component of dairy farms in the Netherlands. Although farm size was the least among groups, Group 4 recorded the highest mean costs on manure compared to groups with higher mean herd sizes. Farms in Group 4 should aim at reducing the cost of manure, and adopt measures to increase prices of milk which was lower than the mean of the population. Unfortunately, the price of milk is exogenous and farmers have little or no control over. Approaches such as value addition and contract sales may help farms to gain higher milk price.

4.2.5 Group 5: High TE, intermediate performance on cluster variables
Group 5 had the highest mean milk price and low mean cost of animal health and the lowest Kg FPCM. The mean herd size, pasture size, concentrate costs, machinery and building depreciation and culling rates were intermediate in Group 5. The group had the worst performance on AFC (statistically same as group 4), breeding and control and cost of bedding (HousnC). Group 5 recorded the second-best mean TE. The high mean TE found can be explained by the highest mean milk price received for milk sold. Empirical findings show milk prices and concentrate prices are the most important determinants of dairy profitability due to the undifferentiated nature of the commodity milk, and is positively correlated with TE (Gloy, 2002; Geert, 2011) The Pearson’s correlation matrix in this study shows that milk price and TE are positively correlated. Group 5 reported the highest mean depreciation costs. The high cost of depreciation could be due to farms having installed new farm assets. Another reason could be that assets are not employed efficiently for productive services. Besides, one important explanation to the high depreciation is the low KgFCPM. The low mean milk yield could account also for the high cost of farm debt (COFD, table 5) of this group. The coefficients of determination of the short-term and long-term determinants are 0.243(24.3%) and 0.245(24.5%) respectively. The cost of building/100kgmilk (RealEst, table 5) and age at first calving (AFC, table 5) negatively correlated with TE in this group.

Dairy farms within Group 5 could improve economic performances by improving milk yield, and produce at the optimal health, breeding, and housing costs. An optimal health and breeding cost lowers AFC which is an important source of reducing feed cost (Hoffman et al., 1999). Group 5 should implement feed formulation strategies that are cost effective. Also, farms should improve the pasture feed which is important for correcting low milk yield (Geert, 2011). The Group should implement initiatives to reduce the cost of borrowing per unit milk produced among farms.
This study found various methods that can be applied in a more practical way to benchmark dairy farms (Section 2.2). The DEA-Cluster methodology proposed and applied in this study has been robust and effective for analysing the joint impact of technical and economic factors influencing economic performance of farms. Future studies based on the proposed method can be more precise by incorporating data mining methods namely, artificial neural networks(ANN) and decision tree(DT). The addition of ANN and DT methods to DEA-Cluster model has good potential to enhancing the capacity to develop a sound classifier selection and identification rules for predicting efficiencies of dairy farms. Besides, future studies can consider applying statistical matching combined with DEA analysis. The major drawback is the pre-determination of the relationships among cluster variables which must be as precise as possible. In view of the limitations of DEA (section 2.3.4), further investigations based on the current study should apply proximity based target selection and stratification methods when conducting the efficiency analysis.

The regression analysis showed that in 4 out of the 5 groups, the long-term variables had greater influence in predicting the TE among farms than the short-term variables (Table 7). The impact of long-term determinants on TE is complex. Zhu et al., (2012) observed a lower impact of the long-term determinant (subsidy) on TE than short-term variable (farm income). Kumbhakar and Lien (2010) observed a positive correlation between the long-term variable(subsidy) with TE. Therefore, as dairy farms strive to improve on TE level, attention should be given to improving on their performances on the long-term determinants as there is greater potential for improving the economic performance among dairy.
Table 7: Summary results of group by group regression analysis

<table>
<thead>
<tr>
<th></th>
<th>Group 1</th>
<th>Group 2</th>
<th>Group 3</th>
<th>Group 4</th>
<th>Group 5</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Long-term Determinants</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Constant($\beta_0$)</td>
<td>0.777***</td>
<td>0.974***</td>
<td>0.784***</td>
<td>0.996***</td>
<td>1.095***</td>
</tr>
<tr>
<td>$R^2_{\text{Short-term}}$</td>
<td>0.300</td>
<td>0.069</td>
<td>0.191</td>
<td>0.054</td>
<td>0.243</td>
</tr>
<tr>
<td>$R^2_{\text{Long-term}}$</td>
<td>0.233</td>
<td>0.135</td>
<td>0.197</td>
<td>0.432</td>
<td>0.245</td>
</tr>
<tr>
<td>VetCost</td>
<td>-0.055*</td>
<td>0.038</td>
<td>-0.056*</td>
<td>0.031**</td>
<td>-0.025</td>
</tr>
<tr>
<td>B&amp;CtrlCost</td>
<td>0.067</td>
<td>-0.014</td>
<td>0.043</td>
<td>0.009</td>
<td>0.078</td>
</tr>
<tr>
<td>HousnCost</td>
<td>-0.013</td>
<td>-0.012</td>
<td>0.007</td>
<td>0.008</td>
<td>0.045</td>
</tr>
<tr>
<td>ManureCost</td>
<td>-0.004</td>
<td>0.004</td>
<td>-0.020</td>
<td>0.006</td>
<td>-0.016</td>
</tr>
<tr>
<td>Dep.M&amp;B</td>
<td>-0.005</td>
<td>-0.002</td>
<td>-0.002</td>
<td>-0.003***</td>
<td>-0.007</td>
</tr>
<tr>
<td>COFD</td>
<td>0.004</td>
<td>0.001</td>
<td>0.008</td>
<td>-0.004</td>
<td>-0.002</td>
</tr>
<tr>
<td>CMPrice</td>
<td>0.000</td>
<td>-0.003</td>
<td>0.001</td>
<td>0.000</td>
<td>-0.002</td>
</tr>
<tr>
<td>RealEst</td>
<td>0.008</td>
<td>-0.008*</td>
<td>-0.014</td>
<td>9.263E-5</td>
<td>-0.022**</td>
</tr>
<tr>
<td><strong>Short-term Determinants</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Constant($\beta_0$)</td>
<td>0.966***</td>
<td>0.837***</td>
<td>0.892**</td>
<td>1.111***</td>
<td>1.217***</td>
</tr>
<tr>
<td>KgFCPM</td>
<td>2.363E-6</td>
<td>1.629E-5</td>
<td>1.755E-5</td>
<td>-3.610E-6</td>
<td>1.262E-5</td>
</tr>
<tr>
<td>KgConc.</td>
<td>-0.007</td>
<td>-0.003</td>
<td>-0.002</td>
<td>0.002</td>
<td>0.016</td>
</tr>
<tr>
<td>Pasture Size</td>
<td>-0.002</td>
<td>0.001</td>
<td>0.001*</td>
<td>0.000</td>
<td>0.000</td>
</tr>
<tr>
<td>Cows</td>
<td>0.000</td>
<td>0.000</td>
<td>0.000**</td>
<td>4.519E-6</td>
<td>-3.465E-5</td>
</tr>
<tr>
<td>CULL</td>
<td>-0.001</td>
<td>0.001</td>
<td>0.001</td>
<td>0.000</td>
<td>0.000</td>
</tr>
<tr>
<td>AFC</td>
<td>-4.483E-5</td>
<td>0.000</td>
<td>0.000</td>
<td>0.000</td>
<td>-0.001**</td>
</tr>
</tbody>
</table>

1\text{VetCost}=Veterinary costs; B&CtrlCost= Breeding and Control cost/100kg milk; HousnCost = Cost of animal bedding/100kg milk; Manure cost=cost to dispose manure off farm/100kg of milk; Dep.M&B= cost of machinery and building depreciation/100kg milk; COFD=Cost of farm debt/100kg milk; CMPrice= critical milk price/100kg milk; RealEst= cost of building/100kg milk produced; KgFCPM= amount of fat content protein milk produce; KgConc. = kg of concentrate feed/100kg milk; Pasture size=is the size of pasture land in ha; Cow= number of cows per herd; CULL=Culling rate of dairy cows; AFC= age of first calving. *significant at $P < 0.1$; **Significant at $P < 0.05$; ***Significant at $P < 0.01$. 

43
Chapter 5: Conclusion

Efficiency analysis (DEA) and cluster analysis have been combined to determine distinct profiles of farm performances among Dutch dairy farms. Combining DEA with multivariable cluster analysis enhanced the discriminatory power of efficiency analysis. Farms were successfully sorted into groups using multivariable clustering without a priori outcomes. Farms were characterized based on the means of economic and the technical variables of the groups. Firstly, the position of dairy farms in the input-output space in reference to the TE of the farms were determined using DEA. Secondly, there was no preconceived judgement about the relationship between other variables and the TE. Thirdly, based on the results from the DEA-Cluster model, advice for improving economic performance was suggested. From the 5 Groups, farms were characterised as high, intermediate, and low performing. Few variables (VetCost, concentrate, and AFC, table, 6) showed segregation across clusters. We found distinct profiles among the Groups; Groups with high TE and high overall performance on cluster variables, high TE and relatively low overall performance, intermediate TE and intermediate performances, low TE and low overall performance, low TE and intermediate performance on cluster variables were of much interest in this study. The outcomes of this study can be a useful starting point for further enquiry about the relationship between the technical variables and economic variables of dairy farms. Based on the cluster analysis and the Pearson correlation matrix, further useful explanations can be found by investigating in-depth the possible interaction effects of dairy productive and economic variables on farm performance to validate the results per cluster. Moreover, the variables included in the analysis represent important productive and economic performances of a dairy farm, other variables can be added for large scale study.

The regression analysis showed that greater gains in economic performance can be achieved by improving on the long-term variables. Analysis based on the proposed model can be enhanced by incorporating artificial neural network, and or decision tree. When analysing efficiency, we suggest the proximity based target selection or stratification can add to enhancing the accuracy of the analysis. The insight gained from this study can is a useful material for extensionists when educating farmers on dairy performance. However, care must be taken when interpreting the results to a single dairy farmer since the analysis were based on the means of farms within clusters and farm performances are very farm-specific.
References


Poppe, K., & van Asseldonk, M. (2015). EIP-AGRI focus group benchmarking of farm productivity and farm performance: How can farmers and advisors use benchmarking data and process to improve productivity and sustainability performance? *EIP-AGRI.*


