

See discussions, stats, and author profiles for this publication at: <https://www.researchgate.net/publication/343125840>

A cluster analysis of variables essential for climate change adaptation of smallholder dairy farmers of Nandi County, Kenya

Article in African Journal of Agricultural Research · July 2020

DOI: 10.5897/AJAR2020.14965

CITATIONS

0

READS

107

4 authors:



Jesse Owino

Kenya Forest Research Institute

19 PUBLICATIONS 48 CITATIONS

SEE PROFILE



Daniel Olago

University of Nairobi

139 PUBLICATIONS 2,709 CITATIONS

SEE PROFILE



Shem Wandiga

University of Nairobi

44 PUBLICATIONS 130 CITATIONS

SEE PROFILE



Oghaiki Asaah Ndambi

Wageningen University & Research

69 PUBLICATIONS 601 CITATIONS

SEE PROFILE

Some of the authors of this publication are also working on these related projects:



Dairy input service delivery systems [View project](#)



DairyBISS - Dairy Business Information Supply & Support [View project](#)

Full Length Research Paper

A cluster analysis of variables essential for climate change adaptation of smallholder dairy farmers of Nandi County, Kenya

Jesse O. Owino^{1,2*}, Daniel Olago¹, Shem O. Wandiga¹ and Asaah Ndambi³

¹Institute for Climate Change Adaptation (ICCA), College of Biological and Physical Sciences, University of Nairobi, P. O. Box 30197, 00100 Nairobi, Kenya.

²Rift Valley Ecoregion Research Program, Kenya Forestry Research Institute, P. O. Box 382-20203, Londiani, Kenya.

³Wageningen Livestock Research, Wageningen University and Research P. O. Box 338, 6700 AB Wageningen, Netherlands.

Received 11 May, 2020; Accepted 6 July, 2020

Smallholder dairy farmers occupy high potential areas of Kenya and are a source of manure, crops and milk. There is need to use other means of characterising smallholder dairy farmers as they mostly practice mixed farming. The objective of this paper is to use cluster analysis method to characterize the smallholder dairy farmers with added farmer and activity data variables. Clusters of 336 farmers in this study were derived using 28 key variables. This paper demonstrates how to conduct farmer assessments for climate change adaptation activities, climate smart technologies implementation using knowledge of key farmer variables and their distribution in the smallholder dairy farmers of Nandi County, Kenya. This paper demonstrates the importance of integrating agricultural information for smallholder dairy farmers to machine models to characterize the groups and observe the natural groupings. This allows for policy managers to know the key characteristics and how to use them in policy implementation especially in designing climate change adaptation programs factoring education and training of farmers as demonstrated in this paper that they are practicing many activities on their farms.

Key words: Cluster analysis, smallholder dairy farmers, farm utilisation, climate change adaptation.

INTRODUCTION

Sub-Saharan Africa (SSA) has the fastest growth in agriculture and the greatest level of agricultural imports compared to other global regions (Livingston et al.,

2011). This growth follows huge demand for food and thus the importance of smallholder agriculture to the food security is cemented for this SSA (Bellarby et al., 2014).

*Corresponding author. E-mail: owinojesse@gmail.com.

Author(s) agree that this article remain permanently open access under the terms of the [Creative Commons Attribution License 4.0 International License](https://creativecommons.org/licenses/by/4.0/)

Smallholder farmers play a key role in development in Africa especially in recent years (Hammond et al., 2015; Livingston et al., 2011; Salami et al., 2010). Smallholder farmers are always characterised in many studies based on land size (Bebe et al., 2002; Herrero et al., 2014; van Averbeké and Mohamed, 2006).

Climate change adverse impacts are heavily felt by the smallholder farmers such as changes on climate variables of rainfall and temperature (Altieri and Koochafkan, 2008; Koochafkan et al., 2012). There exists great diversity in smallholder farmers despite the key variable used for characterizing smallholder farmers, that is farm size (Bebe et al., 2002; Brandt et al., 2018). This diversity is observed as smallholders farmers do mix livestock and crop production and such diversity is seen in smallholder dairy farmers in SSA (Bebe et al., 2002; Oborn et al., 2017; Staal et al., 2002). One of the main reasons climate change affects smallholder farmers is the extra mining of nutrients on their farms, which leads to lower production and also higher susceptibility to climate change effects (Bationo et al., 2004; Castellanos-Navarrete et al., 2015; Rufino et al., 2007).

Smallholder dairy farmers' adaptation to climate change would need the characterization of these farmers with more variables other than farm size as the only limiting factor (Auburger et al., 2015; Herrero et al., 2014; van Averbeké and Mohamed, 2006; Vrieling et al., 2011). There is need for the characterisation of smallholder dairy farmers with additional variables such as land use sizes of the farms, labour, type of livestock housing and other key variables that define smallholders' enterprise (Nyambo et al., 2019; Staal et al., 2002; Waitthaka et al., 2007; Zake et al., 2010). These characterisation should use robust unsupervised models and integrated systems research to show natural groupings of smallholder dairy farmers to allow the policy makers to promote climate change adaptation practices (Nyambo et al., 2019; Oborn et al., 2017; Thompson, 2016). The lack of a characterisation of smallholder farmer with robust inclusion of key variables based on land survey limits options for policy makers as farm size reduces over time (Bebe et al., 2003). This paper seeks to use cluster analysis to characterize smallholder dairy farmers of Nandi County Kenya.

MATERIALS AND METHODS

Study area

The field study was conducted within Nandi County, Kenya (0.565°N, 34.736°E, 0.565°N, 35.437°E, 35.437°E, 0.118°S, 34.736°E, 0.118°S). Mean annual temperatures ranges from 18-22°C, with temperatures at lower elevations (< 1400 m) going as high as 26°C. Altitude ranges from approximately 600 m a.s.l. in the South to over 2200 m a.s.l. in the North east of the county. The highlands are recognized for their high agricultural potential (GOK,

2015; Mudavadi et al., 2001). Nevertheless, livestock and crop farming is mainly subsistence with average land sizes of approximately 4.5 ha per household. Dairy production is common throughout the county, with tea as a major cash crop, and maize as the primary staple crop (GOK, 2015).

Experimental design

Site classification

Agro-ecological zones (AEZ) were identified on the basis of altitude, rainfall, temperature and predominant land use (GOK, 2015). This resulted in three major AEZs: lower highland 1 (LH1: 1900-2400 m a.s.l., area of 934.3 km², high seasonal variation in rainfall and thus having distinct long and short rains, main crops tea and maize). Lower highland 2 (LH2: 1400-1900 m a.s.l., area 1100.7 km², low seasonal variation in rainfall characterized by bimodal rainfall – November-January as the short rains and May-July as the long rains, main crops tea and maize); upper midlands (UM: 1200-1400 m a.s.l. and an area of 364.7 km² with high seasonal variation in rainfall, main crops Sugarcane and maize). A participatory mapping exercise was conducted using experts' knowledge of personnel from the International Livestock Research Institute (ILRI) and Nandi county government to predict whether there were differences in dairy production systems across the AEZs. Thirty-six sampling points were generated with QGIS based on nearness to road infrastructure and masked away from forested areas with the assumption of no households on roads or in forests. The sampling points were kept away from forests since forests are legally gazetted by section 64 (1) of the Forest Conservation and Management Act, number 34 of 2016, prohibiting grazing (Republic of Kenya, 2016). The number of sampling points assigned to each of the three AEZ was weighted by the area of each individual AEZ (cluster), resulting in 15 sites each being located in LH1 and LH2 and six sites being located in the UM (Figure 1). At each of the 36 randomly selected points nine farmers were interviewed to generate a sample size of 336 households.

Household surveys and questionnaire

The household surveys were done using a questionnaire tool customized from the Integrated Modelling Platform for mixed Animal Crop systems (IMPACTlite) (Rufino et al., 2013). IMPACTlite was modified from IMPACT (Herrero et al., 2007) to collect household-level data, which was detailed enough to capture within-site variability on key farm performance and livelihood indicators. It was initially developed to encourage data sharing through standard protocols, and allowing tools to be linked to facilitate evaluations of various farming systems (Rufino et al., 2013). The household questionnaire was completed through face-to-face interviews with the household head using the Open Data Kit (ODK) platform (ODK, 2017). In case of absence of the household head, the most senior member available or the household member responsible for the farm was interviewed. A determination of the primary income categories (crop, dairy, poultry and others) as well as farm size and other farmer demographic data on literacy, age and gender were collected. From this, four animal confinement systems were defined: 'fence only (F)', 'fence and floor (FF)', 'fence and roof (FR)' and 'fence, roof and floor (FRF)'. The animal confinements formed the base for manure management systems so as to be able to relate the confinement with the manure management systems in use.

Manure management systems were classified according to Table

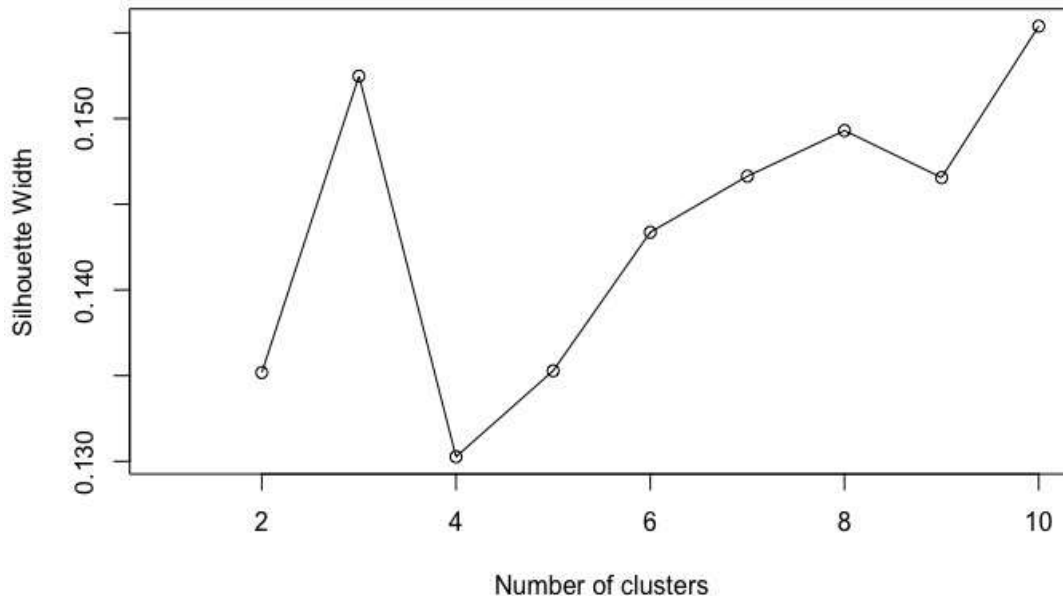


Figure 1. Results of silhouette plot analysis for the Nandi County dataset showing the optimal four clusters.

10.18 provided in the guidelines for GHG emission estimates (IPCC, 2006) and were characterized based on the state of manure being deposited, the location where the manure is stored as well as the duration of storage. The state of manure was either defined as 'fresh'- period less than 24 h from excretion, 'dry' - period more than 24 h from excretion, or 'bioslurry' – that is from biodigesters or farms with pit or lagoon, where liquid manure is collected. Manure handling and storage was characterized as: 'heap for composting', 'pit for fresh or dry manure', 'heap of either fresh or dry manure' and 'pit/lagoon for slurry'. The duration of manure storage before utilization on farm is a proxy for manure quality. Therefore, classification of manure storage was done according to three periods: less than 1 month equalling good quality, 3-4 months meaning reduced quality, and greater than four months equalling least manure quality.

Cluster analysis

Unsupervised learning algorithm was used for cluster analysis. This algorithms was K-means (Chibanda et al., 2009; Nyambo et al., 2019). In the analysis, the number of groups (K) represented how many farm typologies (clusters) could be defined for each dataset. The number of clusters that best represented the data was determined using the Elbow method (where a bend or elbow in a graph showing decline of within cluster sum of squares differences as the number of clusters increases provides the best solution) (Nyambo et al., 2019) The elbow method examines the percentage of variance explained by the clustering as a function of the number of clusters k (Kingrani et al., 2017; Syakur et al., 2018). The K-means algorithm has been widely used in non-hierarchical clustering and characterizing smallholder dairy farms (Kingrani et al., 2017; Nyambo et al., 2019; Tittonell et al., 2010). The algorithm uses Euclidean distance measures to estimate weights of data records. The algorithm is presented as Equation 1, with a segment of the Euclidean distance as in Equation 1.

$$J = \sum_{j=1}^k \sum_{i=1}^n \|x_i^j - c_j\|^2 \quad (1)$$

where $\|x_i^j - c_j\|$ computes the Euclidean distance as in Equation 1; k = number of clusters, n =number of observations, j =minimum number of clusters, i = minimum number of observations, x_i =Euclidean vector for any i th observation, and c_j =cluster centre for any j th cluster. Production cluster outputted from the clustering algorithm was validated in three ways: (1) assessment of cluster robustness, (2) comparison of the cluster membership reallocation (differential allocation of households to clusters for training and testing datasets), and (3) evaluation of the proportion of variation explained by the clusters.

Feature selection

The top 28 features synthesised from literature on smallholder dairy farmers were tabulated (Table 1). These variables have been known to influence productivity in smallholder dairy farming based on experts' domain knowledge. These features and their amounts were Boolean, Discrete and continuous and derived from household survey of 336 smallholder dairy farmers in Nandi County. These variables would be used to identify 'natural groupings' of these 28 features to derive the number and type of clusters (Chibanda et al., 2009; Nyambo et al., 2019; Syakur et al., 2018). This was done by minimising the squared Euclidean distance within a decreasing number of clusters containing an increasing number of positively related variables and using Base R Package (RStudio V 1.1.442) within which dendrogram and plot showing optimal number of clusters using k means was generated (Chibanda et al., 2009). Each of the variables used in clustering was described as percentages (gender, education level, income

Table 1. Variables as features used in cluster analysis of smallholder dairy farmers of Nandi County, Kenya.

S/N	Feature name	Type	Range
1	Agro-ecological Zone	Discrete	1 (LH1), 2 (LH2)-3(UM)
2	Relationship to household head	Discrete	1 (Head), 2 (Spouse), 3 (Child)-4 (Others)
3	Gender	Boolean	1 (Male)- 2 (Female)
4	Age of farmer	Continuous	18-79
5	Education level	Discrete	1(No formal and illiterate), 2(No formal but literate),3 (Primary School), 4 (High school), 5 (College)-6 (University)
6	Total labour available	Continuous	1-13
7	Total time labour required	Continuous	0-26
8	Household area acreage (Hectares)	Continuous	0.05-15
9	Cash crop area acreage (Hectares)	Continuous	0-100
10	Horticulture acreage (Hectares)	Continuous	0-8
11	Grazing area acreage (Hectares)	Continuous	0-48.5
12	Acreage for trees (Hectares)	Continuous	0-180
13	Total farm acreage (Hectares)	Continuous	0.2-210
14	Main Income category	Discrete	1 (Crop), 2 (Pigs), 3 (Poultry), 4(Beef), 5(Dairy)-6 (Other)
15	Total number of dairy cattle	Continuous	1-223
16	Total number of beef cattle	Continuous	0-2
17	Total number of sheep/goats	Continuous	0-42
18	Total number of poultry	Continuous	0-740
19	Total number of other livestock	Continuous	0-8
20	Number of months milking	Continuous	0-12
21	Use of milk commercial/noncommercial	Discrete	1 (Use 100%), 2(Sell >25%), 3 (Sell 25%-75%) -4 (Sell >25%)
22	Number of hours livestock in confinement	Continuous	0-24
23	Number of hours livestock out confinement	Continuous	0-24
24	Livestock confinement system	Discrete	1(Fence Only), 2(Fence and Roof), 3 (Fence and Floor)-4 (Fence, Roof and Floor)
25	is water used in cleaning	Boolean	0(No) -1(Yes)
26	Is bedding removed during cleaning	Boolean	0(No) -1(Yes)
27	Number of manure management systems in use	Discrete	0-4
28	Do the farmers feed concentrates to livestock	Boolean	0(No) -1(Yes)

category) and means (age, acreage under grazing, total acreage, household labour numbers, household dairy numbers and number of manure management systems in use per household). Prediction accuracies were obtained by developing the clustering model in a training dataset

(70% of all records) and the resulting model reapplied to a testing dataset (remaining 30%) (Nyambo et al., 2019). Rank analysis using the spearman correlation coefficient was used to evaluate the level of features reallocation between clusters. Hierarchical clustering was applied and it

works in a bottom-up manner. That is, each variable object is initially considered as a single-element cluster (leaf). At each step of the algorithm, the two clusters that are the most similar are combined into a new bigger cluster (nodes). This procedure is iterated until all points are

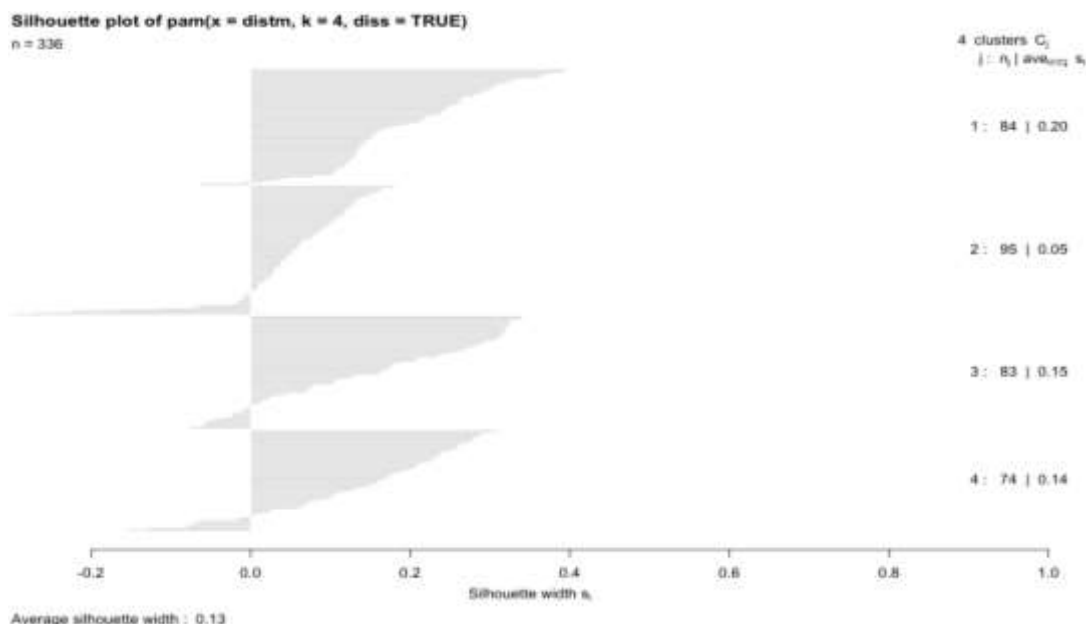


Figure 2. Cluster analysis results showing the four cluster of farmers in Nandi County derived from combining the variables (in column numbers).

member of just one single big cluster (root). Quality control was achieved by using cross clustering in a partial clustering algorithm that combines the Complete Linkage algorithms and Ward's minimum variance providing automatic estimation of a suitable number of clusters and identification of outlier elements (Tellaroli et al., 2016).

RESULTS AND DISCUSSION

Clustering

Based on the Elbow method, a four-cluster solution was found to be optimal for Nandi County dataset and was fitted in the clustering model (Figures 1 and 2). This was confirmed even by quality control using cross clustering and outliers. Cluster analysis was used to classify the smallholder farmers and examine how key variables of acreage for grazing, total acreage, education level, number of dairy cattle in the households and manure management affect their labour practices and major income categories and also classify them (Table 2). This agrees with observations from other studies on smallholders where such variables were enumerated (Nyambo et al., 2019). The study found that there are four classes split by gender and major income categories.

These clusters when using discrete variables showed focus areas such as the low education level of farmers and they were majorly male dominated. These clusters had total low acreage, as well as areas available for grazing, the farmers had less labour and high dairy

livestock numbers. This finding agrees with Chibanda et al. (2009), Tiftonell et al. (2010) and van Averbek and Mohamed (2006) whose studies found that cluster analysis gave the key variables of note to define the components for practice change. There has been studies showing how farmer education can improve various farm practices and subsequently make climate change adaptation (Ausden, 2014; Boswell et al., 2010; Waithaka et al., 2007; Zake et al., 2010). The basic aim of cluster analysis is to find the "natural groupings", if any, of a set of individuals (cases or variables). This was an objective of this study and also agreed with other studies where the advantages of using cluster analysis were extolled (Adeyemo et al., 2019; Chibanda et al., 2009; Kwale, 2013). This study found that after running the analysis, the variables main income category, labour numbers, dairy livestock populations and grazing area acreage were key driving forces for the smallholder farmers. This leads to four clusters based on gender and education levels with differences occurring on the quantities for the other variables. Thus, this study's four clusters were the natural groupings of the smallholder dairy farmers.

CONCLUSION AND RECOMMENDATION

There are many ways to use cluster analysis. The kind of cluster analysis utilised in this study is how to form similar sets of variables. The purpose of this analysis in this study is, therefore, to draw inferences about natural

Table 2. Variables and their statistics in the four clusters of smallholder dairy farmers of Nandi County, Kenya.

Cluster number 1	Variables from Table 1	1	2	3	4	5	6	7	8	9	10																										
n=84		Mean	2	Mean	2	Mean	Male	Mean	44.94	Mean	High School	Mean	2.81	Mean	7.107	Mean	0.4387	Mean	3.206	Mean	0.3821																
Cluster number 2	Variables from Table 1	1	2	3	4	5		6	7	8	9	10																									
n=95		Mean	2	Mean	2	Mean	Male	Mean	43.94	Mean	High School	Mean	3.011	Mean	7.611	Mean	0.4079	Mean	5.158	Mean	0.3789																
Cluster number 3	Variables from Table 1	1	2	3	4	5		6	7	8	9	10																									
n=83		Mean	1	Mean	2	Mean	Female	Mean	42.05	Mean	Primary School	Mean	2.47	Mean	6.952	Mean	0.8482	Mean	3.128	Mean	0.462																
Cluster number 4	Variables from Table 1	1	2	3	4	5		6	7	8	9	10																									
n=74		Mean	2	Mean	2	Mean	Male	Mean	43.96	Mean	High School	Mean	3.203	Mean	9.297	Mean	0.6372	Mean	2.172	Mean	0.2014																
Cluster number 1	Variables from Table 1		11		12				13		14		15		16		17		18		19																
n=84			Mean		3.885		Mean		2.774		Mean		10.685		Mean		8.512		Mean		0.05952		Mean		3.119		Mean		13.51		Mean		0.1786				
Cluster number 2	Variables from Table 1		11		12				13		14		15		16		17		18		19																
n=95			Mean		1.829		Mean		0.3516		Mean		8.127		Mean		11.41		Mean		0.01053		Mean		2.526		Mean		27.92		Mean		0.08511				
Cluster number 3	Variables from Table 1		11		12				13		14		15		16		17		18		19																
n=83			Mean		3.814		Mean		2.558		Mean		9.106		Mean		8.892		Mean		0.03614		Mean		2.422		Mean		19.24		Mean		0.1084				
Cluster number 4	Variables from Table 1		11		12				13		14		15		16		17		18		19																
n=74			Mean		1.852		Mean		0.2905		Mean		5.153		Mean		4.324		Mean		0.01351		Mean		1.541		Mean		9.689		Mean		0.1081				
Cluster number 1	Variables from Table 1		20		21				22		23		24		25		26		27		28																
n=84			Mean		9.893		Mean		Sold >75%		Mean		22.33		Mean		1.667		Mean		Fence Only		Mean		0.05952		Mean		0.0119		Mean		1		Mean		Yes
Cluster number 2	Variables from Table 1		20		21				22		23		24		25		26		27		28																
n=95			Mean		8.2		Mean		Sold <25%		Mean		21.27		Mean		2.474		Mean		Fence Only		Mean		0.09474		Mean		0.02105		Mean		1		Mean		Yes
Cluster number 3	Variables from Table 1		20		21				22		23		24		25		26		27		28																
n=83			Mean		10.73		Mean		Sold >75%		Mean		13.96		Mean		10.33		Mean		Fence Only		Mean		0.04819		Mean		0.0241		Mean		1		Mean		Yes
Cluster number 4	Variables from Table 1		20		21				22		23		24		25		26		27		28																
n=74			Mean		9.135		Mean		Sold 25-75%		Mean		12.97		Mean		11.03		Mean		Fence Only		Mean		0		Mean		0		Mean		0		Mean		No

groupings of smallholder dairy farmers and the nature of the key variables used in these groupings. Hierarchical clustering was appropriate for and could also be applied to qualitative

variables (Kwale, 2013; Nyambo et al., 2019). The major result of this study was that with as many as 28 variables the cluster analysis revealed only four distinct natural groupings. In this study,

cluster analysis is used to test the proposition that there are simple natural groupings of smallholder dairy farmers and they are mirrors with realisation that the analysis yielded distinct groupings

(Chibanda et al., 2009; Condliffe et al., 2008). The study recommends there should be future research to give a detailed characterisation of other areas and types of farmers using cluster analysis to compare the counties and also scenarios if scaled to the region.

CONFLICT OF INTERESTS

The authors have not declared any conflict of interests.

REFERENCES

- Adeyemo T, Amaza P, Okoruwa V, Akinyosoye V, Salman K, Abass A (2019). Determinants of intensity of biomass utilization: Evidence from cassava smallholders in Nigeria. *Sustainability (Switzerland)* 11:9.
- Altieri M, Koohafkan P (2008). *Enduring Farms: Climate Change, Smallholders and Traditional Farming Communities*. In Third World Network Penang, Malaysia. http://www.fao.org/nr/water/docs/enduring_farms.pdf
- Auburger S, Wustholz R, Petig E, Bahrs E (2015). Biogas digestate and its economic impact on farms and biogas plants according to the upper limit for nitrogen spreading—The case of nutrient-burdened areas in north-west Germany. *AIMS Energy* 3(4):740-759.
- Ausden M (2014). Climate Change Adaptation: Putting Principles into Practice. *Environmental Management* 54(4):685–698.
- Bationo A, Nandwa SM, Kimetu JM, Kinyangi JM, Bado BV (2004). Sustainable intensification of crop – livestock systems through manure management in eastern and western Africa: Lessons learned and emerging research opportunities Africa, pp. 173-198.
- Bebe BO, Udo HMJ, Rowlands GJ, Thorpe W (2003). Smallholder dairy systems in the Kenya highlands: Cattle population dynamics under increasing intensification. *Livestock Production Science* 82(2-3):211-221.
- Bebe BO, Udo HMJ, Thorpe W (2002). Development of smallholder dairy systems in the Kenya highlands. *Outlook on Agriculture* 31(2):113-120.
- Bellarby J, Stirling C, Vetter SH, Kassie M, Kanampiu F, Sonder K, Smith P, Hillier J (2014). Identifying secure and low carbon food production practices: A case study in Kenya and Ethiopia. *Agriculture, Ecosystems and Environment* 197:137-146.
- Boswell MR, Greve AI, Seale TL (2010). An Assessment of the Link Between Greenhouse Gas Emissions Inventories and Climate Action Plans. *Journal of the American Planning Association* 76(4):451-462.
- Brandt P, Hamunyela E, Herold M, de Bruin S, Verbesselt J, Rufino MC (2018). Sustainable intensification of dairy production can reduce forest disturbance in Kenyan montane forests. *Agriculture, Ecosystems and Environment* 265:307-319.
- Castellanos-Navarrete A, Tittonell P, Rufino MC, Giller KE (2015). Feeding, crop residue and manure management for integrated soil fertility management – A case study from Kenya. *Agricultural Systems* 134:24-35.
- Chibanda M, Ortmann GF, Lyne MC (2009). Institutional and governance factors influencing the performance of selected smallholder agricultural cooperatives in KwaZulu-Natal. *Agrekon* 48(3):293–315.
- Condliffe K, Love C, Porter M (2008). *Kenya Coffee: A Cluster Analysis*. Harvard Business School Microeconomics of Competitiveness P. 2. https://www.isc.hbs.edu/resources/courses/moc-course-at-harvard/Documents/pdf/student-projects/Kenya_Coffee_2008.pdf
- GOK (2015). Nandi County Development Profile. Government of Kenya. nandi.go.ke/wp-content/uploads/2015/04/Nandi-County-Profile.doc
- Hammond J, Fraval S, van Etten J, Suchini JG, Mercado L, Pagella T, Frelat R, Lannerstad M, Douchamps S, Teufel N, Valbuena D, van Wijk MT (2015). The rural household multi-indicator survey (RHoMIS) for rapid characterisation of households to inform climate smart agriculture interventions: Description and applications in East Africa and Central America. *Agricultural Systems*, May. <https://doi.org/10.1016/j.agsy.2016.05.003>
- Herrero M, González-Estrada E, Thornton PK, Quirós C, Waitaha MM, Ruiz R, Hoogenboom G (2007). IMPACT: Generic household-level databases and diagnostics tools for integrated crop-livestock systems analysis. *Agricultural Systems* 92(1-3):240-265.
- Herrero M, Thornton PK, Bernués A, Baltenweck I, Vervoort J, van de Steeg J, Makokha S, van Wijk MT, Karanja S, Rufino MC, Staal SJ (2014). Exploring future changes in smallholder farming systems by linking socio-economic scenarios with regional and household models. *Global Environmental Change* 24(1):165-182.
- IPCC (2006). Guidelines for national greenhouse gas inventories. Chapter 10: Emissions from livestock and manure management. http://www.ipcc-nggip.iges.or.jp/public/2006gl/pdf/4_Volume4/V4_10_Ch10_Livestock.pdf
- Kingrani SK, Levene M, Zhang D (2017). Estimating the number of clusters using diversity. *Artificial Intelligence Research* 7(1):15. <https://doi.org/10.5430/air.v7n1p15>
- Koohafkan P, Altieri MA, Holt Gimenez E (2012). Green Agriculture: Foundations for biodiverse, resilient and productive agricultural systems. *International Journal of Agricultural Sustainability* 10(1):61-75.
- Kwale FM (2013). A Critical Review of K Means Text Clustering Algorithms. *International Journal of Advanced Research in Computer Science* 04(09):27-34.
- Livingston G, Schonberger S, Delaney S (2011). Sub-Saharan Africa: The state of smallholders in agriculture. IFAD: International Fund for Agricultural Development: Conference on New Directions for Smallholder Agriculture, pp. 1-36.
- Mudavadi PO, Otieno K, Wanambacha JW, Odenya JO (2001). Smallholder dairy production and marketing in western Kenya: A review of literature. <https://cgspace.cgiar.org/handle/10568/1781>
- Nyambo DG, Luhanga ET, Yonah ZO, Mujibi FDN (2019). Application of Multiple Unsupervised Models to Validate Clusters Robustness in Characterizing Smallholder Dairy Farmers. *Scientific World Journal*, 2019. <https://doi.org/10.1155/2019/1020521>
- Oborn I, Vanlauwe B, Phillips M, Thomas R, Brooijmans W, Atta-Krah K (2017). Sustainable Intensification in Smallholder Agriculture. <https://doi.org/10.4324/9781315618791>
- ODK (2017). Open Data Kit. <https://opendatakit.org/>
- Republic of Kenya (2016). The Forest Conservation and Management Act. Kenya Gazette Supplement. kenyalaw.org/kl/.../Acts/ForestConservationandManagementActNo34of2016.pdf
- Rufino MC, Tittonell P, van Wijk MT, Castellanos-Navarrete A, Delve RJ, de Ridder N, Giller KE (2007). Manure as a key resource within smallholder farming systems: Analysing farm-scale nutrient cycling efficiencies with the NUANCES framework. *Livestock Science* 112(3):273-287. <https://doi.org/10.1016/j.livsci.2007.09.011>
- Rufino MC, Quiros C, Boureima M, Desta S, Douchamps S, Herrero M, Kiplimo J, Lamissa D, Mango J, Moussa AS, Naab J, Ndour Y, Sayula G, Silvestri S, Singh D, Teufel N, Wanyama I (2013). Developing generic tools for characterizing agricultural systems for climate and global change studies (IMPACTlite – phase 2) (Issue January). <https://cgspace.cgiar.org/handle/10568/42065>
- Salami A, Kamara AB, Brixiova Z (2010). Smallholder Agriculture in East Africa: Trends, Constraints and Opportunities. In Working Paper No.105 African Development Bank (Issue April). <https://doi.org/10.1111/j.1467-937X.2007.00447.x>
- Staal SJ, Baltenweck I, Waitaha MM, DeWolff T, Njoroge L (2002). Location and uptake: Integrated household and GIS analysis of technology adoption and land use, with application to smallholder dairy farms in Kenya. *Agricultural Economics* 27(3):295-315. [https://doi.org/10.1016/S0169-5150\(02\)00075-0](https://doi.org/10.1016/S0169-5150(02)00075-0)
- Syakur MA, Khotimah BK, Rochman EMS, Satoto BD (2018). Integration K-Means Clustering Method and Elbow Method for Identification of the Best Customer Profile Cluster. IOP Conference

- Series: Materials Science and Engineering 336:1.
- Tellaroli P, Bazzi M, Donato M, Brazzale AR, Drăghici S (2016). Cross-Clustering: A Partial Clustering Algorithm with Automatic Estimation of the Number of Clusters. *PLoS One* 11(3):e0152333.
- Thompson D (2016). Restoring Fire to Grasslands: An Overview in Fire Management and Restoration in Rangelands. Proceedings of the 10th International Rangeland Congress, July, pp. 454-455.
- Tittonell P, Muriuki A, Shepherd KD, Mugendi D, Kaizzi KC, Okeyo J, Verchot L, Coe R, Vanlauwe B (2010). The diversity of rural livelihoods and their influence on soil fertility in agricultural systems of East Africa - A typology of smallholder farms. *Agricultural Systems* 103(2):83-97.
- van Averbeke W, Mohamed SS (2006). Smallholder farming styles and development policy in South Africa: The case of Dzindi Irrigation Scheme. *Agricon* 45(2):136-157.
- Vrieling A, Beurs KM, Brown ME (2011). Variability of African farming systems from phenological analysis of NDVI time series. *Climatic Change* 109(3):455-477.
- Waithaka MM, Thornton PK, Shepherd KD, Ndiwa NN (2007). Factors affecting the use of fertilizers and manure by smallholders: The case of Vihiga, western Kenya. *Nutrient Cycling in Agroecosystems* 78(3):211-224.
- Zake J, Tenywa JS, Kabi F (2010). Improvement of manure management for crop production in Central Uganda. *Journal of Sustainable Agriculture* 34(6):595-617.