Agrekon

Agricultural Economics Research, Policy and Practice in Southern Africa

ISSN: (Print) (Online) Journal homepage: www.tandfonline.com/journals/ragr20

Identifying clusters of innovation barriers in farming enterprises: *a K-modes clustering approach*

Yasser Buchana

To cite this article: Yasser Buchana (19 Feb 2025): Identifying clusters of innovation barriers in farming enterprises: *a K-modes clustering approach*, Agrekon, DOI: 10.1080/03031853.2024.2441127

To link to this article: https://doi.org/10.1080/03031853.2024.2441127

Published online: 19 Feb 2025.

AGREKON
https://doi.org/10.1080/03031853.2024.2441127

Identifying clusters of innovation barriers in farming enterprises: *a K-modes clustering approach*

Yasser Buchana

Centre for Science Technology and Innovation Indicators, Human Science Research Council City, Cape Town, South Africa

ABSTRACT
Despite the relevance of agriculture in terms of its contribution to food security and environmental sustainability, empirical evidence on the nature and effects of barriers to innovation, remains very limited. This problem has often led to inconsistent policy design, failing to meet the needs and expectations of farming enterprises. Innovation profiling and segmentation have emerged as important tools for understanding how businesses innovate. Using data from the South African Agricultural Business Innovation Survey (AgriBIS-2016-2018), this study applies the *K*-modes clustering algorithm, to group farming enterprises based on the innovation barriers they face. The analysis identified three distinct clusters with differing innovation profiles. The findings show that businesses in Cluster 2 recognised institutional barriers such as Lack of Government Support (65.4%) and Stringent Agricultural Policies (93.6%) as key barriers to their innovations, while Cluster 3 businesses highlighted environmental factors as critical impediments to their innovation. Cluster 1 businesses faced a diverse range of resource-related barriers. This study addresses a key knowledge gap and increases our understanding of the barriers to agricultural innovation. The findings have important implications for innovation policy instruments and strategies aimed at promoting agricultural innovation, sustainability, productivity and resilience in the face of barriers to innovation.

ARTICLE HISTORY
Received 22 November 2022
Accepted 6 December 2024

KEYWORDS
Cluster analysis; *K*-modes algorithm; machine learning; innovation barriers; agriculture

1. Introduction

In recent years, innovation profiling and segmentation have become an essential methods in academic and policy circles for analysing the innovation behaviour of firms (Arundel and Hollanders 2004; Peneder 2010; Simmie 2004). When firms innovate, they are often faced with a wide variety of challenges or impediments. These are more pronounced in the agricultural sector, especially, since these have experienced a variety of obstacles that have hampered their innovation efforts (Buchana and Sithole 2022; Dahabieh, Bröring, and Maine 2018). These obstacles include, among other factors, resource (financial and human), institutional and regulatory, as well as environmental factors (climate change, droughts, floods, etc.).

In an effort to address some of the challenges to innovation faced by farming enterprises, policy makers have often applied a blanket approach to policy design which, in some circumstances, has failed to meet the needs and requirements of agricultural businesses (Hassan 2010; Pretty et al. 2001). Some of these policies have not produced the desired effects due to a lack of understanding

CONTACT Yasser Buchana ybuchana@gmail.com; ybuchana@hsrc.ac.za

of the characteristics and profiles of innovation in farming enterprises and, in particular, the barriers to innovation in these firms. This is reflected in weak innovation performance in farming enterprises, which has remained much lower when compared to other sectors of the economy, especially those in developing nations (Spielman, 2005). Innovation remains nonetheless a critical vehicle for addressing the sector's challenges, notably those of poor productivity, provided that a conducive framework is in place to support it (Buchana and Sithole 2022; Ortmann 2005; Rambe and Khaola 2022). This can help ensure the agricultural sector's long-term viability and sustainability (Ulvenblad et al. 2018).

Despite their use, simple aggregate indices of innovation-active enterprises that consider a specific barrier to innovation as "very important" provide little relevant information for policy makers. To better understand the structure and complexity of barriers to innovation in the agricultural sector, businesses should be profiled based on the barriers they face. Alleviating or at least minimising barriers to innovation may stimulate more innovations in the sector, thereby increasing the number of businesses actively engaged in innovation activities (Popkova 2022; Silva et al. 2023).

While agriculture is vital for food security and environmental sustainability, empirical evidence on the nature and effects of barriers to innovation, as well as how these barriers affect businesses' ability to engage in innovation activities, remains very limited (Sivertsson and Tell 2015). There are multiple reasons for this, but two main factors stand out. First, there are few studies on barriers to innovation in general, not only in agriculture, and those that have been conducted have generally focused on sectors of the economy such as manufacturing and services. Second, the limited research in this field has not focused on characterising firms based on the barriers they face using analytical techniques such as cluster analysis. As a result, little is known about the profiles of farming enterprises based on the barriers to innovation. As such, the primary research question guiding this study is formulated as follows: *How can the barriers to innovation in farming enterprises be effectively profiled and clustered to inform targeted innovation support policies?*

Using data from the South African Agricultural Business Innovation Survey (AgriBIS-2016-2018), this study applied the *K*-modes algorithm to analyse the clusters of farming enterprises that have experienced a wide range of innovation barriers. *K*-modes cluster analysis is a type of unsupervised learning clustering technique designed to identify latent patterns within a set of unlabelled data (Khan and Ahmad 2013). The primary objective of the *K*-modes clustering approach is to divide a dataset with several attributes into distinct groups, often referred to as clusters. The goal is to ensure that the data objects inside a cluster have a higher degree of similarity to one another, as determined by a similarity measure, compared to objects in other clusters (Huang 1998). As such, cluster analysis using this approach enables the discovery of hidden trends in the data, as well as allows for the identification of innovation profiles and behaviours (Roelandt and den Hertog 1999).

This study fills a significant knowledge gap and advances our understanding of the barriers to innovation in the agriculture sector. The research also has significant implications for innovation policy instruments that aim to promote innovation, productivity, sustainability, and resilience by removing barriers to innovation.

2. Literature review

There are three major streams of research on the phenomenon of barriers to innovation. The first stream of research has mainly looked at obstacles to innovation primarily from the standpoint of factors that influence how barriers to innovation are perceived by businesses (D'Este et al. 2012; Iammarino, Sanna-Randaccio, and Savona 2009). The second and more common has mostly focused on the effects that impede such obstacles (Coad, Pellegrino, and Savona 2016; Pellegrino 2018). However, the underlying denominator among these different approaches is that they have all mostly focused their analyses on innovation performance and intensity instead of diving deeper into the impact of these barriers on economic performance, which are mediated by innovation (Popkova 2022; Silva et al. 2023). More recently, the perception of (mostly financial) barriers to innovation and their impediment effects on the decisions of businesses to engage in innovation activities has gained the attention of innovation scholars (see for example Cinar, Trott, and Simms 2019). This growing attention extends to the degree and propensity of businesses to engage in innovation (Baldwin and Lin 2002; Galia and Legros 2004; Iammarino, Sanna-Randaccio, and Savona 2009).

Nonetheless, existing but very limited literature on the phenomenon of barriers to innovation has mostly been dedicated to sectors of the economy such as manufacturing and services (Savignac 2008). In addition, this literature has mostly concentrated on factors that influence successful innovation outcomes in those sectors. Few studies have focused on factors that impede innovation (Mu and Wang 2022). More specifically, the majority of this rare research has focused on theoretical and empirical implications of financial barriers on the innovation behaviour of businesses (see Sandberg and Aarikka-Stenroos 2014). While financial resources are necessary for the successful development of new products and processes in businesses, recent research has shown that there are other key factors that can potentially have a substantial negative impact on a firm's innovation efforts. Skill shortage, a lack of demand for specific products and services as well as a lack of relevant understanding of technology and markets are some of the factors highlighted in the literature (see Blanchard et al. 2013; Mohnen et al. 2008).

A third stream of research, the literature on innovation management, has paid special attention to the factors that influence innovation failure. This is in addition to factors that influence successful innovation outcomes. This line of research has mostly examined how different types of businesses are susceptible to various types of innovation barriers (Segarra-Blasco, Garcia-Quevedo, and Teruel-Carrizosa 2008). One of the most prominent debates in this field of research is the distinguishing innovation characteristics and capabilities of large, medium and small businesses. According to the innovation management literature, large and more established businesses have the required profiles, innovation capabilities and large budgets for developing radical innovations. However, they seldom do; instead, they tend to develop more incremental innovations (Leifer et al. 2000; Oliveira et al. 2019; Vercher, Bosworth, and Esparcia 2023). On the other hand, smaller and emerging businesses who are rarely equipped for developing radical innovations often do produce radical innovations. Nevertheless, because of their heterogeneity and innovation capacity, all these businesses face distinct barriers to innovation and these constraints have varied influences on their productivity (Galli et al. 2020; Saunila 2020).

Meanwhile, existing debates within the innovation literature have consistently ignored the factors that may impede or limit the capacity of farming enterprises to engage in innovation in order to improve their productivity performance (Eastwood et al. 2023). In comparison to other sectors of the economy, the literature on barriers to innovation in agriculture is less substantial, although it is slowly emerging. Much like other sectors of the economy, businesses in the agricultural sector also face similar, but more specialised, types of barriers to innovations (Ulvenblad et al. 2018). Some studies have used the resource-based theory perspective approach to categorise barriers to innovation into two types of barriers; (a) internal barriers and (b) external barriers (De Faria, Noseleit, and Los 2020; Madrid-Guijarro, Garcia, and Van Auken 2009; Radicic 2021). Examples of internal barriers include: lack of access to farming land; lack of access to farming equipment; lack of access to water; and lack of access to agricultural loans and subsidies. All of these may be classified as barriers to innovation in the manner stated above (Avellaneda-Rivera 2019; Ulvenblad et al. 2018). Examples of external barriers include: weather and climate change; increased global competition; and burdensome agricultural policies and regulations (Ulvenblad et al. 2018). In other words, they are barriers that are beyond the control of farming enterprises.

In conclusion, as can be seen from the discussion in this review of literature, existing debates on barriers to innovation have a number of significant theoretical and methodological implications. More importantly, these implications are more pronounced in the agricultural sector, where policy makers and innovation scholars have struggled to understand the nature and magnitude of the effects of barriers to innovation on the productivity of farming enterprises.

3. Theoretical foundations

3.1 Conceptualising innovation clusters within farming enterprises

The study of clusters has emerged as a key phenomenon in various fields, including computer science, bioinformatics, statistics and even marketing, helping researchers to identify and discover new patterns in data (Dalmaijer, Nord, and Astle 2022; Duran and Odell 2013; Kaufman and Rousseeuw 2009; Punj and Stewart 1983).

A "cluster" can be defined broadly as a comparable grouping of objects. The fundamental ideas behind the concept of cluster are similarity in characteristics, closeness and density (Feser and Luger 2003; Hollenstein 2003).

In farming enterprises, clusters can significantly influence innovation through the dense networks of information flow, shared resources and collaborative efforts they facilitate (Preissl and Solimene 2003; Mazur et al. 2016). However, innovation clusters also face distinct barriers, such as limited access to new technologies or markets and regulatory constraints (Pihkala, Ylinenpaa, and Vesalainen 2002; Varman and Chakrabarti 2011). Understanding the dynamics and barriers within these clusters is necessary for identifying targeted strategies to promote innovation in the agricultural sector.

In the field of computer science, bio-informatics and data mining, one of the extensively used approaches to cluster large data sets efficiently is the *k*-means algorithm (e.g., Aghakhani, Qabaja, and Alhajj 2018; Cui et al. 2014; Nerurkar et al. 2018). The *k*-means algorithm is an algorithm that groups data based on similar characteristics by calculating the mean of the data points, hence the "means" in "*k*-means" (Pena, Lozano, and Larranaga 1999). However, the *k*-means algorithm was designed to only handle numerical data and therefore does not work well with categorical data. This is the main drawback of the *k*-means algorithm, despite its efficiency. To overcome this limitation, it has been common practice to first transform nominal variables into discreet variables, each of which represents the existence or absence of a category (Huang 1998). This approach also has its own drawbacks. As a consequence, a dataset with many nominal variables may generate a large number of discrete variables, which would significantly increase the computing complexity and memory requirements for the clustering, making this approach highly inefficient (Huang 1998).

Realising the limitations of the *k*-means algorithm, Huang (1998) designed the *K*-modes clustering algorithm by extending the existing *k*-means approach to accommodate categorical data. When creating clusters, Huang's *K*-modes algorithm substitutes modes for means and uses the matching dissimilarity measure for categorical variables as well as a frequency-based approach to update the modes.

4. Methods and dataset

4.1 Analysing innovation barriers using the K*-modes clustering approach*

The choice of *K*-modes clustering as our analytical method was informed by the nature of our data and the theoretical constructs of our study. Given that the barriers to innovation in farming enterprises are often categorical (e.g., regulatory constraints, access to technology, market limitations), *K*-modes clustering presents a methodologically sound approach to categorising these barriers into distinct clusters. This method is particularly suited for analysing categorical data, as it uses modes instead of means for clustering, thereby preserving the categorical nature of our data without necessitating transformation into numerical values.

4.2 Description of how the K*-modes algorithm works*

The *K*-modes clustering algorithm works by first determining the number of clusters. It then divides the data into "*k*" distinct segments or partitions by assigning each observation to the closest cluster, then re-partitions the data multiple times until the optimum clustering quality is achieved.

The steps for the *K*-modes algorithm are given by Table 1 below as follows:

Table 1. Description of steps of *K*-modes algorithms.

Steps	Description
Step 1	Initialise *K* modes by predetermining the desired number of modes.
Step 2	Based on a simple dissimilarity score, assign the items to the nearest cluster. After each allocation, update each cluster mode.
Step 3	Check each item's dissimilarity value against the mode once all the items have been assigned to a cluster. If it turns out that an item's closest mode is in a different cluster, relocate the item to that cluster and update both clusters' modes.
Step 4	Repeat step 3 until none of the items change to another cluster.

4.3 Implementation of K-modes clustering algorithm in this study

This study used a dataset of farming enterprises, where each row represents a particular firm, and each column represents a specific barrier to innovation. The dissimilarity measure, which is the Hamming distance or simple matching dissimilarity, allowed us to calculate and compare how different or similar the firms were in terms of the barriers to innovation they experienced.

Each firm was assumed to be a unique entity, and the barriers to be the attributes that vary between the different firms. A value of "1" in a particular column indicates that the firm considered that specific barrier as highly important for their innovation, while a "0" means they did not consider it important.

In this case, the algorithm allowed us to calculate and compare how different or similar these firms were, based on their attributes. The more differences (mismatches), the greater the dissimilarity, which then implied that these firms had distinct perspectives on innovation barriers. Fewer differences indicated that they have similar perceptions on these barriers, which suggests a higher degree of similarity between the firms in terms of their innovation priorities.

Suppose A and B are two firms in our farming enterprises innovation dataset, and m is the total number of innovation barriers in the dataset. Equation (1) below describes the simple matching dissimilarity measure between two firms (***A*** and ***B***) based on the barriers they experienced.

$$d(A, B) = \sum_{i=1}^{m} \delta(A_i, B_i) \quad (1)$$

Where,

- *d(A, B)* is the Hamming distance or dissimilarity measure between firms A and B.
- A_i *and* B_i represent the values (1 or 0) of the *i*-th barrier for firms A and B, respectively.
- *m* is the total number of innovation barriers in the dataset.

4.4 Data sources and variables

This study used data from the South African baseline Agricultural Business Innovation Survey covering the period 2016–2018 (Agri-BIS 2016–2018). The Agri-BIS 2016–2018 was based on the guidelines of the Organisation for Economic Co-operation and Development's (OECD) Oslo Manual (OECD/Eurostat 2005). The survey used the methodological recommendations for the Community Innovation Survey (CIS) of the European Union (EU) countries, as provided by Eurostat, the Statistical Office of the European Commission. The survey focused on ascertaining how agricultural businesses innovate. The core questions asked about the businesses' product, process, organisational and marketing innovations. The survey also asked questions about the different innovation activities and outcomes. To determine the barriers to innovation, the survey incorporated additional questions on the factors that impede agricultural innovation. These questions asked businesses about the different factors that they considered highly important during the reference period 2016–2018.

The population for the baseline South African AgriBIS 2016-2018 survey was defined as agricultural enterprises operating within the agriculture, forestry, and fisheries subsectors, classified under Standard Industrial Codes (SIC) 11, 12, and 13. Statistics South Africa (Stats SA) constructed the sampling frame using these SIC codes, ensuring representation across the subsectors. The sample for the baseline survey comprised 1 690 agricultural enterprises, with initial subsector allocations of 1514 for agriculture, 95 for forestry, and 81 for fisheries.

Following data cleaning, 364 invalid entries (e.g., duplicates or untraceable entities) were removed and the final sample was reduced to 1326 enterprises. Respondents were categorised into four size groups, i.e. large, medium, small, and very small, based on turnover thresholds adjusted from the 2003 Department of Trade and Industry (DTI) guidelines. Annual turnover thresholds were adjusted to reflect contemporary business scales. Large enterprises were defined as those with a turnover exceeding R40 million, while medium enterprises reported turnover between R24 million and R40 million. Small enterprises fell within the turnover range of R4 million to R24 million, and very small enterprises reported turnover of up to R4 million.

Following the analysis of the sampling frame, in consultation between Stats SA and the Centre for Science, Technology and Innovation Indicators (CeSTII), these thresholds were derived by applying a factor of eight to the original boundaries set by the Department of Trade and Industry in 2003. This ensured the categorisation aligned with the economic realities of agricultural enterprises and facilitated comparability with official classifications, such as those in the 2017 Agricultural Census.

Data collection was conducted via digital tools, including REDCap for online questionnaire administration using Everlytic email software for survey dispatch and supplemented by Adobe pdf forms available in English and Afrikaans. Monitoring systems tracked survey participation to enable efficient follow-ups. In a difficult business climate, 303 businesses responded to the survey over a short and intensive fieldwork period of three months in 2019. On this basis, the survey achieved an overall response rate of 22%.

A follow up non-response survey targeting 15% of non-respondents was conducted to mitigate bias, as recommended by the OECD (2005). This non-response survey covered 117 businesses and achieved a response rate of 74.3%. Adjusted probability weights, reflecting invalid enterprises and non-response corrections, allowed extrapolation to the population of agricultural enterprises. Respondents' profiles included a mix of racial groups, geographical locations, turnover bands, and enterprise types. The survey's representativeness was further validated through adherence to the South African Statistical Quality Assessment Framework (SASQAF) and alignment with official enterprise categories from the Stats SA Agricultural Census (2017). This methodological framework ensured that findings reasonably informed conclusions about innovation activities across South Africa's agricultural enterprises.

The survey covered three main subsectors of commercial agricultural businesses at the higher level of classification: the agriculture subsector (e.g., crop producers, wineries, livestock and poultry), the forestry subsector and the fisheries subsector.

The sample obtained from Stats SA for the baseline AgriBIS 2016–2018 did not include data on the ownership structure of agricultural enterprises, and the survey itself not collect data on the race of respondents, ownership and location. Therefore these aspects could not be included in the analysis.

However, based on turnover, innovation activity, and employment size, the majority of firms were classified as either Large (42.2%) or Medium-sized (40.3%), with a smaller share identified as Small (12.9%) or Very Small (4.6%). After further cleaning the dataset to focus exclusively on firms involved in animal and crop farming (SIC 11), a similar distribution was observed, with Large enterprises making up 42.4%, Medium-sized firms 40.6%, Small enterprises 13.3%, and Very Small enterprises 3.7%.

However, this study used observations from the agriculture subsector, that is, businesses that were in SIC 11 of the South African agricultural sector. These were businesses that were involved in animal and crop farming. The rationale of excluding the forestry and fisheries subsectors is that they included very few observations and would have potentially caused misclassification. Furthermore, the analysis only included innovation-active businesses. This data set contained 214 units. After removing units that had missing values, the final data set included 168 observations. Missing values can lead to misclassification; therefore, it was necessary to remove them (Table 2).

Table 2. Variables selected for the cluster analysis.

Variable name	Description
Lack_of_Access_finance	If a farming enterprise responded as having experienced lack of access to finance as a barrier. A dummy variable was constructed with: 1 = High, Medium 0 = Low, Not relevant
Lack_of_Access_land	If a farming enterprise responded as having experienced lack of access to land as a barrier. A dummy variable was constructed with: 1 = High, Medium 0 = Low, Not relevant
Lack_of_Access_water	If a farming enterprise responded as having experienced lack of access to water as a barrier. A dummy variable was constructed with: 1 = High, Medium 0 = Low, Not relevant
Lack_of_Access_agro-chemicals	If a farming enterprise responded as having experienced lack of access to agrochemicals as a barrier. A dummy variable was constructed with: 1 = High, Medium 0 = Low, Not relevant
Lack_of_Access_labour	If a farming enterprise responded as having experienced lack of access to labour as a barrier. A dummy variable was constructed with: 1 = High, Medium 0 = Low, Not relevant
Weather_Climate_change	If a farming enterprise responded as having experienced weather / climate change as a highly important barrier. A dummy variable was constructed with: 1 = High, Medium 0 = Low, Not relevant

Python programming language was used to implement the *K*-modes machine learning algorithm on the dataset using the Scikit-Learn and Pandas library in Python. The Pandas library is the most popular library for manipulating and analysing data using data frames in Python. It includes a suite of tools that makes it easier to clean and pre-process data. The Scikit-learn library, on the other hand, provides a collection of Python tools for machine learning which comprises classification, regression, clustering, dimensionality reduction, and model selection methods. This study applied these two libraries for data pre-processing and machine learning (classification, clustering, association rules and visualisation).

5. Analysis and findings

In this study, the clustering model for *K*-modes was implemented with cost function in getting the minimal distance for the intra cluster distance. The Elbow method is a commonly used technique to determine the optimal number of clusters, and it involves evaluating the cost function across a range of *K* values. The idea is to look for an "elbow point" where the cost function begins to level off, which then indicates a diminishing return on increasing the number of clusters (Khan and Ahmad 2013). The Elbow method was used to determine the optimum number of clusters (*K* value for *K*-modes), which the model was set to choose between two and four clusters (i.e., $k = 2$, 3 and 4). Figure 1 shows that the elbow shape is detected when number of clusters suggested was $K = 3$ using cost function value.

By aggregating comparable clusters and distinguishing dissimilar ones, *K*-modes clustering was able to categorise the profiles of agricultural enterprises according to the different characteristics of the barriers they experienced. This allowed us to analyse how barriers to innovations affect the different clusters of farming enterprises.

Figure 2 shows the distribution of farming enterprises that were affected by barriers in the three clusters. The bulk of enterprises were located in the second cluster (46% of enterprises), while the first cluster had the second-highest concentration of businesses (30%) and the third cluster included the remaining businesses (24%).

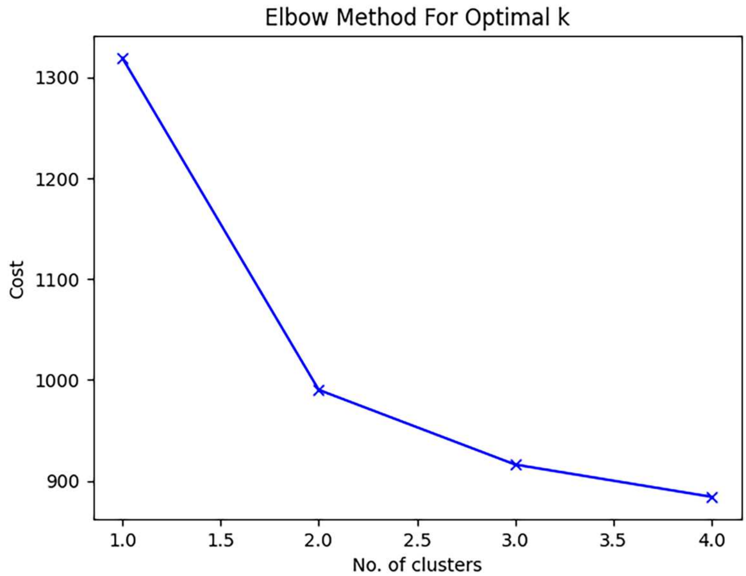

Figure 1. The Elbow method for number of clusters.

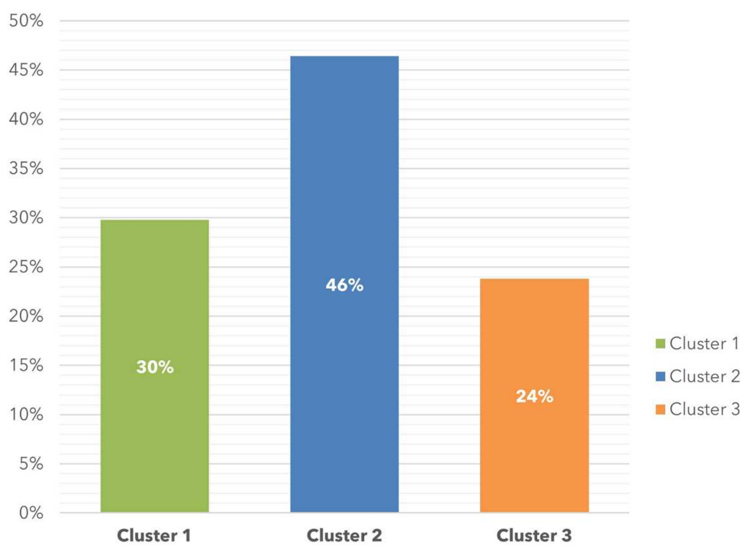

Figure 2. Distribution of farming enterprises in the clusters.

The next step was to determine the characteristics of each cluster based on the barriers experienced. Cluster analysis revealed hidden trends in the data, allowing for the discovery of innovation profiles and behaviours.

For enhanced analysis and comparison, a radar graph was used to visualise the different clusters of the barriers to innovation that affect businesses.

As depicted by the radar graph in Figure 3, the majority of farming enterprises in the second cluster consider all barriers as highly important. In contrast to the first cluster, a greater proportion of enterprises in the first cluster consider poor agricultural policies, lack of access to community support, weather and climate change, lack of government support, lack of access to labour, and

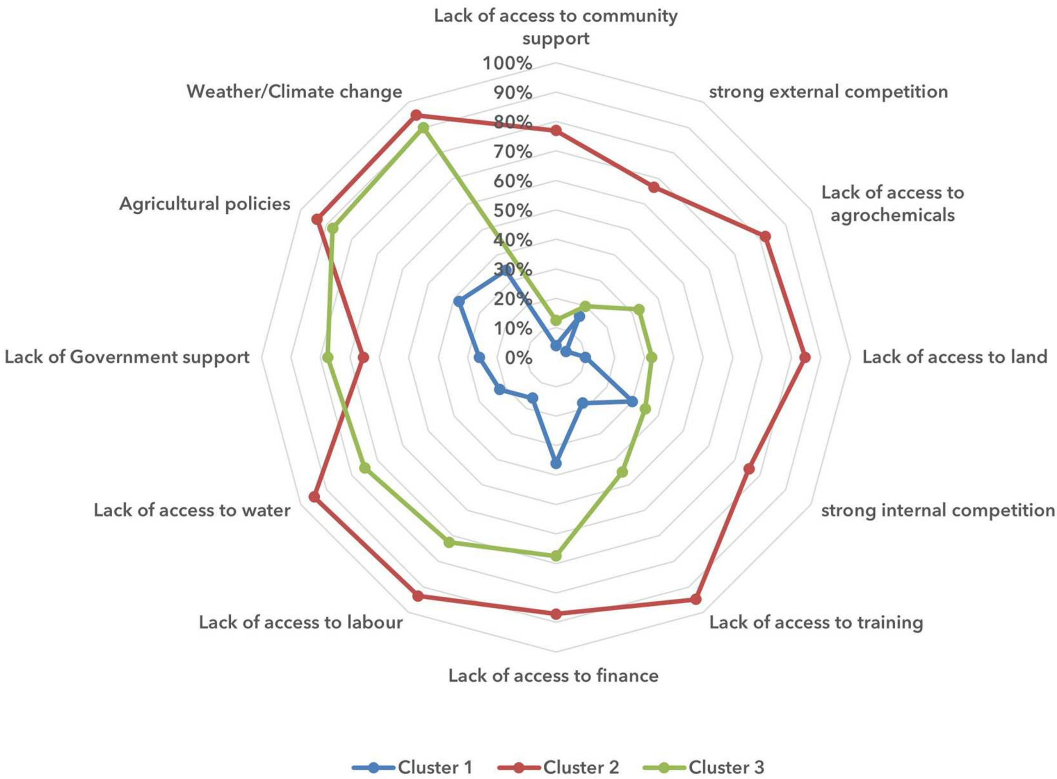

Figure 3. Clusters and how they are affected by different barriers.

lack of access to finance as major impediments to innovation. In the third cluster, a lower proportion of enterprises perceive these barriers to be highly important.

The analysis of the three clusters revealed both commonalities and distinctiveness in their approaches to addressing barriers to innovation in the agricultural sector. One common characteristic that unites these clusters is the shared recognition of the critical role played by *Lack of access to finance*. Previous research has widely acknowledged that financial resources play a significant role in enabling innovation, and that sustained innovation is dependent on sufficient access to financial resources and investments in innovation activities.

All three clusters were reportedly affected by Weather/Climate change. This shared focus by farming enterprises in the three clusters shows how aware these enterprises are with respect to how environmental conditions and climate change affect farming. This is especially significant in South Africa, where weather conditions have had a big effect on farming in the last few years.

The clusters may be differentiated based on their unique characteristics, which in turn influence their own strategies and priorities. For example, Cluster 2 stood out for being the cluster with the most businesses, but also for being the most affected by a wide range of barriers to innovation. In contrast, Cluster 1 was mostly marked by a balanced recognition of both institutional and environmental challenges. Finally, most businesses in Cluster 3 were particularly focused on resource factors and climate resilience.

Although the three clusters showed a few common qualities and with respect to the barriers to innovation, their unique attributes revealed a diverse array of priorities and goals within agricultural innovation landscape in South Africa.

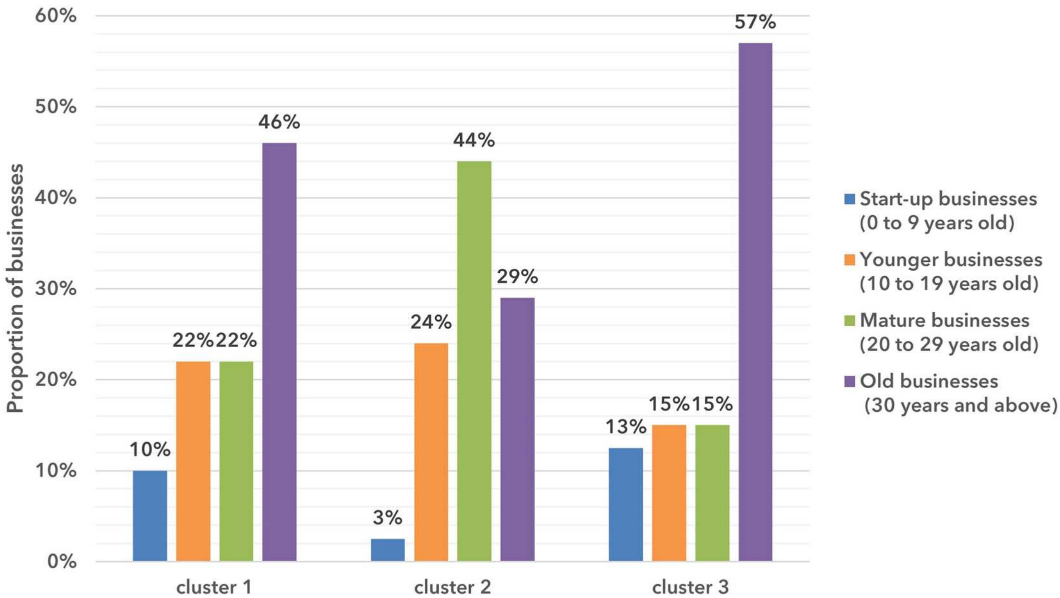

Figure 4. Age group distribution.

6. Profiling of farming enterprises

6.1 Age distribution of enterprises in the clusters

The age distribution of farming enterprises is an important factor to consider since the different age groups of firms were affected differently by innovation barriers. Younger farming enterprises typically find it difficult to overcome such barriers in the same way that older enterprises do, while older and more mature enterprises usually have more financial and non-financial resources, such as skill and experience, to deal with innovation barriers.

As illustrated by Figure 4, the farming enterprises were grouped into four distinct age categories namely: Start-up enterprises, defined as those less than 10 years old, fell into age-group 1. Age-group 2 contained younger enterprises, defined as being between the ages of 10 and 19. In age-group 3, enterprises were between the ages of 20 and 29. Finally, older enterprises which were aged 30 and above fell under age-group 4.

In terms of the three clusters, Cluster 1 emerged as a cluster characterised by the presence of well-established and mature farming enterprises, with a substantial 46% of enterprises falling into the older age group (30 years and above).

In contrast, Cluster 2 contained a substantial proportion of mature farming enterprises (44%), indicating a concentration of enterprises that have demonstrated stability and resilience over time. Meanwhile, Cluster 3 stood out as the cluster predominantly composed of long-standing and old farming enterprises, with a remarkable 57% of enterprises categorised in this age group.

The diverse age profiles within the three clusters highlights the effectiveness of the *K*-modes algorithm in grouping farming enterprises with similar experiences related to innovation barriers and age. Each cluster had a distinct mixture of firms from different age groups, which reflects the complex characteristics of farming enterprises in these clusters.

6.2 Number of innovation activities

Innovation activities can be a useful measure of how much effort businesses invest into their innovation capabilities. Figure 5 shows the analysis of the number of innovation activities per firm in each cluster.

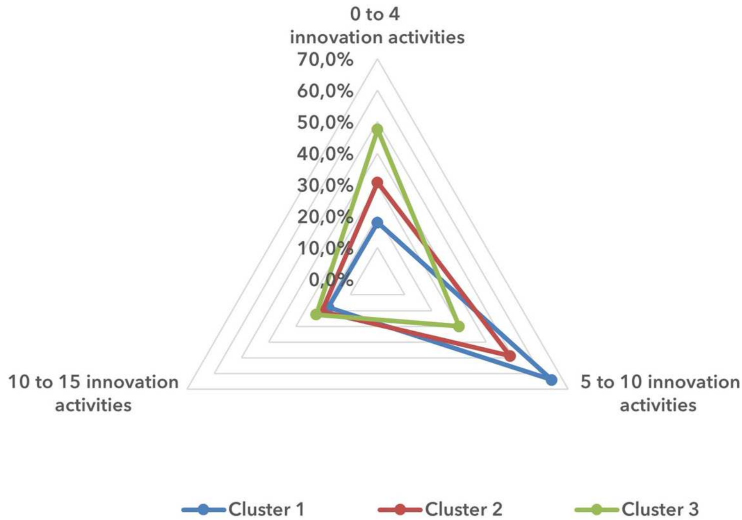

Figure 5. Number of innovation activities groups.

Three categories were used to organise the farming enterprises in terms of the number of innovation activities. Farming enterprises with zero to four innovation activities fell into the first group. The second category included enterprises with five to 10 innovation activities, while the third group included enterprises with 10 or more innovation activities.

The results show that the majority of the farming enterprises in Cluster 1 and Cluster 2 had been engaged in five to 10 innovation activities, while the majority of the enterprises in the third cluster had between zero and five innovation activities.

Cluster 2 was primarily distinguished by a more balanced distribution of farming enterprises engaged in a variety of innovation activities, with 30.8% in 0–5 innovation activities, 48.7% in 5–10 activities and 20.5% in 10–15 activities. Furthermore, this balance remains consistent throughout several age groups. This finding suggests that enterprises in Cluster 2 showed a more flexible and adaptive innovation approach irrespective of their age. In contrast, Cluster 1, which had more enterprises that had between 5–10 innovation activities (64%), may be indicative of a concentrated effort to develop expertise in a particular area of innovation. This specialisation may have been substantially influenced by the age and expertise of farming enterprises in this cluster.

6.3 Resource barriers

The analysis revealed that Cluster 2 stood out as a cluster in which farming enterprises overwhelmingly considered these resource-related barriers to be highly important, as shown in Figure 6. The percentages were notably high across all four barriers, with 87.2% emphasising the lack of access to finance, 84.6% focusing on land-related challenges, 94.8% highlighting water access issues and 76.9% underscoring the significance of community support access barriers. This suggests that Cluster 2 enterprises placed a strong emphasis on addressing these resource-related challenges as a priority.

On the other hand, Cluster 1 had a more balanced approach, where no single barrier dominated in terms of importance. While 36% of enterprises recognised the lack of access to finance as a significant challenge, other barriers such as access to land (10%), water (22%) and community support (4%) were also considered important, but to a much lesser extent. This balance implies that enterprises in Cluster 1 were more diverse in their perceptions about the resource barriers.

In contrast, Cluster 3 had a different profile to the first and second clusters. It had an intriguing profile with varying degrees of importance attributed to the four resource factors. A significant percentage of farming enterprises in Cluster 3 (68%) view the lack of access to finance as a significant

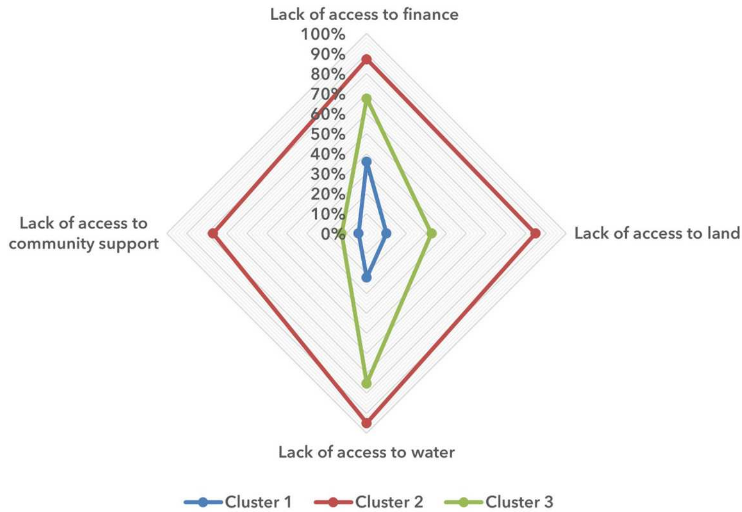

Figure 6. Resource barriers.

challenge. For example, access to water (75%) was considered highly important, possibly indicating water resource management as a primary concern. Interestingly, while land access (33%) and community support (13%) were perceived as important barriers to innovation, they received relatively lower priority compared to access to finance and water.

6.4 Knowledge barriers

Lack of access to labour, lack of access to training/skills, and lack of access to agrochemicals were grouped together into knowledge factors.

As shown in Figure 7, Cluster 2 notably stood out as a cluster where enterprises overwhelmingly prioritised knowledge-related barriers to innovation. The percentages were relatively high across all

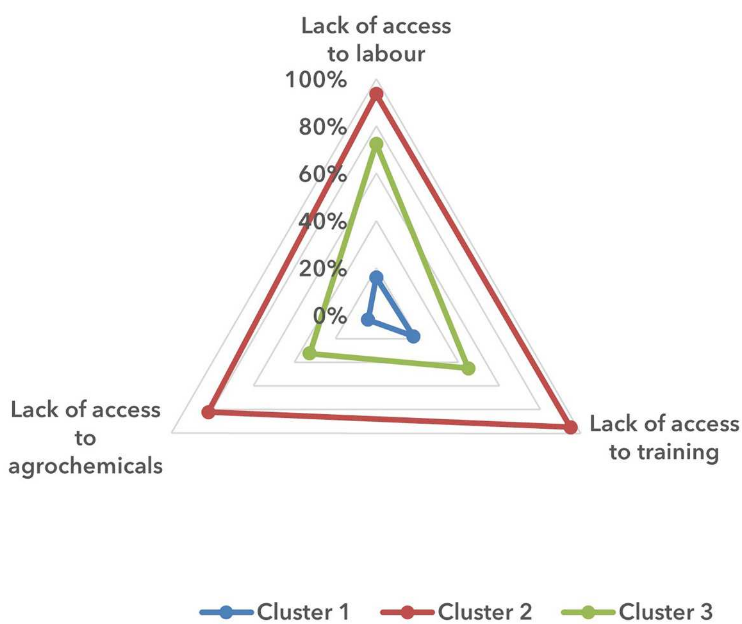

Figure 7. Knowledge barriers.

three knowledge barriers, with 93.6% of businesses highlighting the lack of access to labour, 94.9% emphasising access to training as an important barrier, and 82.1% highlighting access to agrochemicals as an important barrier to their innovation. In contrast, Cluster 3 showed a more balanced approach, where no single barrier dominated in terms of importance. While 72.5% of businesses view the lack of access to labour as a significant challenge, 45% of businesses in the third cluster considered access to training as highly important and 32.5% considered access to agrochemicals as an important barrier.

Meanwhile, Cluster 1 had a different profile altogether with distinct importance among the three knowledge factors. While the lack of access to training (18%) and labour (16%) were considered important, these barriers received a relatively lower degree of importance compared to access to agrochemicals (4%).

6.5 Market factors, institutional and environmental barriers

Figure 8 shows the proportions of firms in the three different clusters based on their perceived importance of market, institutional and environmental barriers to innovation. The analysis shows diverse innovation landscapes within different clusters of farming enterprises. While Cluster 3 placed a strong emphasis on environmental factors (90.0%), Cluster 2 adopted a more comprehensive approach, and Cluster 1 had mixed priorities to address market, institutional and environmental barriers. More specifically, farming enterprises in Cluster 2 placed a high emphasis on institutional barriers like Lack of Government Support (65.4%) and Agricultural Policies (93.6%). This cluster also recognised the significance of market competition, with 75.6% highlighting strong internal competition and 66.7% identifying strong external competition as highly important barriers. Farming enterprises in Cluster 2 also showed a strong concern for environmental factors with approximately 95% of enterprises citing weather and climate change as highly important barriers to innovation.

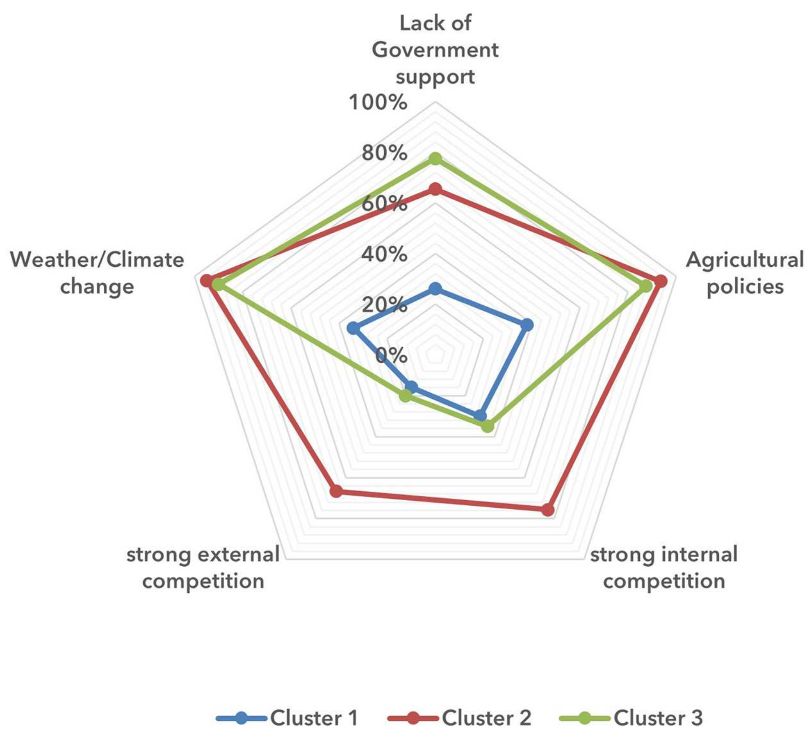

Figure 8. Market, institutional and environmental barriers.

On the other hand, Cluster 1 had a unique profile, with most farming enterprises having mixed priorities with respect to market, institutional and environmental barriers to innovation as illustrated in Figure 8. Previous research has shown that government support, through subsidies, or tax incentives, are considered important mechanisms that facilitate innovation in businesses. This means that institutional factors such as government support and policies must be conducive for innovation to take place. While farming enterprises in Cluster 1 recognised the importance of institutional barriers like agricultural policies (38%) and lack of government support (26%), enterprises in this cluster also acknowledged market competition, particularly strong internal competition (30%). However, less than 40% farming enterprises in this cluster placed significant importance on Weather/Climate Change (34.0%). This mix of priorities indicates that farming enterprises in Cluster 1 adopt a varied approach to address various challenges and adapt to diverse barriers to innovation.

7. Discussion

Previous studies have shown that innovation plays an essential role in addressing issues of poor productivity and ensuring the long-term sustainability of the agricultural sector. While farming enterprises often grapple with a wide range of barriers to their innovation efforts, understanding and alleviating these barriers has recently emerged as a critical step in promoting innovation within the sector. Consequently, profiling and segmenting farming enterprises based on their innovation behaviour, especially the barriers to innovation they face, has become an important area of research.

Until now, much of the debate in the innovation literature on barriers has primarily focused mostly on financial barriers to innovation (Mohnen et al. 2008; Savignac 2008). However, more recent debates have pointed to the existence of other, equally significant barriers that influence the capacity of businesses to innovate beyond financial ones (D'Este et al. 2012; Radicic 2021). These include, for example, knowledge factors, institutional factors and market factors as well as environmental barriers. These barriers, when combined, can impact the ability of businesses to innovate (Sandberg and Aarikka-Stenroos 2014). While financial barriers have traditionally been the focus of attention for scholars and policy makers, this study explores other highly significant barriers to innovation faced by farming enterprises.

Using the *K*-modes clustering algorithm, the analysis presented in this study revealed the intricate profiles of farming enterprises based on the barriers they face, thereby providing an in-depth understanding of the barriers to innovation ecosystem. The *K*-modes clustering approach was chosen for clustering because of its proficiency in handling complex data patterns and its efficiency in dealing with large-scale categorical datasets. The analysis identified three distinct clusters of farming enterprises, each with its own unique set of barriers and priorities.

Table 3 provides a summary overview of each cluster's innovation profile based on the dominant barriers, age profile and level of engagement in innovation activities. The table summarises the key characteristics of each cluster and allows for a quick reference to understand their distinct attributes.

Cluster 1 was predominantly characterised by a low prevalence of innovation barriers, with most farming enterprises reporting low percentages for various barriers. This suggests that these farming enterprises faced relatively fewer barriers to innovation. The low prevalence of barriers could be attributed to their adaptability and experience, given that this cluster mainly consisted of older and more established farming enterprises, and were actively engaged in 5–10 innovation activities.

Table 3. Summary overview of each cluster's innovation profile.

Cluster	Name	Dominant barriers	Age profile	Innovation activities
Cluster 1	Resource constrained	Resource-related	Predominantly older	Engaged in 5–10 activities
Cluster 2	Multifaceted barriers	Diverse sets of barriers across the board	Mixture of younger and mature businesses	Engaged in 0–5 activities
Cluster 3	Environmentally engaged	Environmental focus	Predominantly older	Engaged in 5–10 activities

Their involvement in high number of innovation activities could have been the reason for fewer barriers effectively.

Conversely, Cluster 2 showed a contrasting picture with a substantial proportion of farming enterprises experiencing a diverse set of innovation barriers, resulting in higher percentages of enterprises experiencing many different types of barrier. This cluster included both a mixture of mature and younger farming enterprises, and their collective emphasis on various barriers also reflected the lower number of innovation activities they engaged in.

Cluster 3 had a mixed pattern, with a high prevalence of certain barriers and a low prevalence of others. This cluster predominantly consisted of well-established, older farming enterprises and their focus on specific barriers, particularly environmental factors, suggests a commitment to sustainability and resilience against environmental challenges.

8. Conclusion

Previously, various cluster analysis techniques have been used to build innovation profiles, discover hidden innovation behaviour amongst businesses (Kaufman and Rousseeuw 2009; Roelandt and den Hertog 1999). For this study, a *K*-modes clustering algorithm using machine learning approaches was applied to analyse the clusters of farming enterprises that have experienced a wide range of innovation barriers. Overall, three distinct clusters were derived, and their characteristics were analysed and discussed. The study showed that farming enterprises have diverse innovation profiles, requiring tailored strategies to promote innovation.

The *K*-modes algorithm served as an important clustering tool for discovering these hidden and varied innovation profiles, which would otherwise have been extremely difficult to find. The findings, particularly the innovation profiles, can help guide policy makers in designing effective and targeted policies that promote innovation and resilience in the agricultural sector. These policies should address the needs of farming enterprises, and alleviate barriers to innovation in specific clusters.

This study shows that the different barriers to innovation faced by farming enterprises, render ‘one-size-fits-all’ policies ineffective. To promote a more innovative agricultural sector, this study advocates for the development and implementation of targeted policy measures, specifically designed to address the distinct barriers identified across different clusters of farming enterprises. This approach necessitates a policy framework that not only recognises the unique challenges faced by each cluster, but also facilitates targeted support mechanisms. Through the adoption of a more segmented approach to policy design, government and industry stakeholders can more effectively catalyse innovation within the agricultural sector, to make sure policies are applicable as well as impactful for all types of farming enterprises.

9. Policy implications

Over the last few years, existing innovation policy instruments have often applied a blanket approach to address innovation barriers faced by businesses. These approaches have mostly proven inadequate given that the agricultural sector is still lagging behind other sectors of the economy with respect to innovation. The blanket approach strategy has mostly failed to meet the needs of farming enterprises in terms of barriers to innovation, which has led to neglected opportunities for facilitating innovation in the agricultural sector, particularly in developing countries. While the distinguishing characteristics of the agricultural sector, such as its weak innovation performance, should be taken into account when developing innovation policies aimed primarily at alleviating barriers to innovation, some fundamental principles should apply to all such public-policy instruments supporting innovation.

Through cluster analysis of farming enterprises based on their barriers to innovation, policy makers, more than ever, should have a good understanding of the need for sound policy targeting and justifications for policy initiatives, as well as a strong foundation for government involvement in supporting innovation across the agricultural sector.

However, some critical considerations must be addressed. For effective policy interventions, it is essential to recognise the distinct characteristics and needs of each cluster. For example, in Cluster 1, policy makers should focus on providing accessible financing options and resource optimisation support to stimulate innovation in this group of firms. Meanwhile, in Cluster 2, inclusive support programmes should be designed to address multifaceted barriers, including governance improvements and knowledge-focused capacity building. Lastly, Cluster 3 calls for policies promoting eco-friendly practices, resource management and climate resilience strategies to align with its commitment to environmental sustainability and innovation. These targeted approaches can effectively address the unique innovation challenges faced by farming enterprises in each cluster within the agricultural sector.

Disclosure statement

No potential conflict of interest was reported by the author(s).

References

Aghakhani, S., A. Qabaja, and R. Alhajj. 2018. Integration of k-means clustering algorithm with network analysis for drug-target interactions network prediction. *International Journal of Data Mining and Bioinformatics* 20, no. 3: 185–212.

Arundel, A., and H. Hollanders. 2004. European innovation scoreboard–exis: An exploratory approach to innovation scoreboards. *European Commission, Brussels*.

Avellaneda-Rivera, L. 2019. Why not? Barriers to eco-innovation in the food sector in Spain. *Research on Open-Innovation Strategies and Eco-Innovation in Agro-Food Industries* 37: 37–52.

Baldwin, J., and Z. Lin. 2002. Impediments to advanced technology adoption for Canadian manufacturers. *Research Policy* 31, no. 1: 1–18.

Blanchard, P., J.-P. Huiban, A. Musolesi, and P. Sevestre. 2013. Where there is a will, there is a way? Assessing the impact of obstacles to innovation. *Industrial and Corporate Change* 22, no. 3: 679–710.

Buchana, Y., and M.M. Sithole. 2022. Towards a conceptual framework for measuring innovation in the agricultural sector in sub-Saharan developing countries. *African Journal of Science, Technology, Innovation and Development* 15, no. 2: 272–282.

Cinar, E., P. Trott, and C. Simms. 2019. A systematic review of barriers to public sector innovation process. *Public Management Review* 21, no. 2: 264–290.

Coad, A., G. Pellegrino, and M. Savona. 2016. Barriers to innovation and firm productivity. *Economics of Innovation and New Technology* 25, no. 3: 321–334.

Cui, X., P. Zhu, X. Yang, K. Li, and C. Ji. 2014. Optimized big data k-means clustering using mapreduce. *The Journal of Supercomputing* 70, no. 3: 1249–1259.

Dahabieh, M.S., S. Bröring, and E. Maine. 2018. Overcoming barriers to innovation in food and agricultural biotechnology. *Trends in Food Science & Technology* 79: 204–213.

Dalmaijer, E.S., C.L. Nord, and D.E. Astle. 2022. Statistical power for cluster analysis. *BMC Bioinformatics* 23, no. 1: 1–28.

De Faria, P., F. Noseleit, and B. Los. 2020. The influence of internal barriers on open innovation. *Industry and Innovation* 27, no. 3: 205–209.

D'Este, P., S. Iammarino, M. Savona, and N. von Tunzelmann. 2012. What hampers innovation? Revealed barriers versus deterring barriers. *Research Policy* 41, no. 2: 482–488.

Duran, B.S., and P.L. Odell. 2013. *Cluster analysis: A survey* (Vol. 100). Springer Science & Business Media. Berlin, Heidelberg, New York.

Eastwood, C., J.A. Turner, A. Romera, D. Selbie, R. Henwood, M. Espig, and M. Wever. 2023. A review of multi-scale barriers to transitioning from digital agriculture to a digital bioeconomy. *CABI Reviews* 2023.

Feser, E.J., and M.I. Luger. 2003. Cluster analysis as a mode of inquiry: Its use in science and technology policymaking in North Carolina. *European Planning Studies* 11, no. 1: 11–24.

Galli, F., P. Prosperi, E. Favilli, S. D'Amico, F. Bartolini, and G. Brunori. 2020. How can policy processes remove barriers to sustainable food systems in Europe? Contributing to a policy framework for agri-food transitions. *Food Policy* 96: 101871.

Galia, F., and D. Legros. 2004. Complementarities between obstacles to innovation: evidence from France. *Research Policy* 33, no. 8: 1185–1199.

Hassan, R.M. 2010. Implications of climate change for agricultural sector performance in Africa: Policy challenges and research agenda. *Journal of African Economies* 19, no. suppl_2: ii77–ii105.

Hollenstein, H. 2003. Innovation modes in the Swiss service sector: A cluster analysis based on firm-level data. *Research Policy* 32, no. 5: 845–863.

Huang, Z. 1998. Extensions to the k-means algorithm for clustering large data sets with categorical values. *Data Mining and Knowledge Discovery* 2, no. 3: 283–304.

Iammarino, S., F. Sanna-Randaccio, and M. Savona. 2009. The perception of obstacles to innovation. foreign multinationals and domestic firms in Italy. *Revue d'économie industrielle* 125: 75–104.

Kaufman, L., and P.J. Rousseeuw. 2009. *Finding groups in data: An introduction to cluster analysis*. John Wiley & Sons. New Jersey.

Khan, S.S., and A. Ahmad. 2013. Cluster center initialization algorithm for K-modes clustering. *Expert Systems with Applications* 40, no. 18: 7444–7456.

Leifer, R., C.M. McDermott, G.C. O'connor, L.S. Peters, M.P. Rice, R.W. Veryzer, et al. 2000. *Radical innovation: How mature companies can outsmart upstarts*. Harvard Business Press. Boston, Massachusetts.

Madrid-Guijarro, A., D. Garcia, and H. Van Auken. 2009. Barriers to innovation among Spanish manufacturing smes. *Journal of Small Business Management* 47, no. 4: 465–488.

Mazur, V.V., K.A. Barmuta, S.S. Demin, E.A. Tikhomirov, and M.A. Bykovskiy. 2016. Innovation clusters: Advantages and disadvantages. *International Journal of Economics and Financial Issues* 6, no. 1: 270–274.

Mohnen, P., F.C. Palm, S.S. Van Der Loeff, and A. Tiwari. 2008. Financial constraints and other obstacles: Are they a threat to innovation activity? *De Economist* 156, no. 2: 201–214.

Mu, R., and H. Wang. 2022. A systematic literature review of open innovation in the public sector: Comparing barriers and governance strategies of digital and non-digital open innovation. *Public Management Review* 24, no. 4: 489–511.

Nerurkar, P., A. Shirke, M. Chandane, and S. Bhirud. 2018. Empirical analysis of data clustering algorithms. *Procedia Computer Science* 125: 770–779.

OECD/Eurostat. 2005. *Oslo manual: Guidelines for collecting and interpreting innovation data*. OECD publishing. Paris.

Oliveira, C.A.O., E.P. Ruffoni, A.C.G. Maçada, and ÂD Padula. 2019. Innovation capabilities in the food processing industry in Brazil. *British Food Journal* 121, no. 11: 2901–2918.

Ortmann, G.F. 2005. Promoting the competitiveness of South African agriculture in a dynamic economic and political environment. *Agrekon* 44, no. 3: 286–320.

Pihkala, T., H. Ylinenpaa, and J. Vesalainen. 2002. Innovation barriers amongst clusters of European SMEs. *International Journal of Entrepreneurship and Innovation Management* 2, no. 6: 520–536.

Pellegrino, G. 2018. Barriers to innovation in young and mature firms. *Journal of Evolutionary Economics* 28, no. 1: 181–206.

Pena, J.M., J.A. Lozano, and P. Larranaga. 1999. An empirical comparison of four initialization methods for the k-means algorithm. *Pattern Recognition Letters* 20, no. 10: 1027–1040.

Peneder, M. 2010. Technological regimes and the variety of innovation behaviour: Creating integrated taxonomies of firms and sectors. *Research Policy* 39, no. 3: 323–334.

Popkova, E.G. 2022. Vertical farms based on hydroponics, deep learning, and AI as smart innovation in agriculture. In Popkova, E.G., and B.S. Bruno (Eds.) (2022), *Smart innovation in agriculture*, 257–262. Singapore: Springer Nature Singapore.

Pretty, J., C. Brett, D. Gee, R. Hine, C. Mason, J. Morison, and T. Dobbs. 2001. Policy challenges and priorities for internalizing the externalities of modern agriculture. *Journal of Environmental Planning and Management* 44, no. 2: 263–283.

Preissl, B., and L. Solimene. 2003. *Dynamics of clusters and innovation* (Vol. 62, p. 527). Heidelberg, Germany: Physica-Verlag.

Punj, G., and D.W. Stewart. 1983. Cluster analysis in marketing research: Review and suggestions for application. *Journal of Marketing Research* 20, no. 2: 134–148.

Radicic, D. 2021. Financial and non-financial barriers to innovation and the degree of radicalness. *Sustainability* 13, no. 4: 2179.

Rambe, P., and P. Khaola. 2022. The impact of innovation on agribusiness competitiveness: The mediating role of technology transfer and productivity. *European Journal of Innovation Management* 25, no. 3: 741–773.

Roelandt, T. J. A., and P. den Hertog. 1999. Cluster analysis and cluster-based policy making: the state of the art. In Cluster Analysis and Cluster-based Policy: New perspectives and Rationale in Innovation Policy. In edited by T. Roelandt and P. den Hertog. Paris: Organisation for Economic Cooperation and Development. 413–427.

Sandberg, B., and L. Aarikka-Stenroos. 2014. What makes it so difficult? A systematic review on barriers to radical innovation. *Industrial Marketing Management* 43, no. 8: 1293–1305.

Saunila, M. 2020. Innovation capability in smes: A systematic review of the literature. *Journal of Innovation & Knowledge* 5, no. 4: 260–265.

Savignac, F. 2008. Impact of financial constraints on innovation: What can be learned from a direct measure? *Economics of Innovation and New Technology* 17, no. 6: 553–569.

Segarra-Blasco, A., J. Garcia-Quevedo, and M. Teruel-Carrizosa. 2008. Barriers to innovation and public policy in Catalonia. *International Entrepreneurship and Management Journal* 4, no. 4: 431–451.
Silva, F.T.D., I.C. Baierle, R.G.D.F. Correa, M.A. Sellitto, F.A.P. Peres, and L.M. Kipper. 2023. Open innovation in agribusiness: Barriers and challenges in the transition to agriculture 4.0. *Sustainability* 15, no. 11: 8562.
Simmie, J. 2004. Innovation and clustering in the globalised international economy. *Urban Studies* 41, no. 5–6: 1095–1112.
Sivertsson, O., and J. Tell. 2015. Barriers to business model innovation in Swedish agriculture. *Sustainability* 7, no. 2: 1957–1969.
Spielman, D.J. 2005. Innovation systems perspectives on developing-country agriculture: A critical review.
Ulvenblad, P., H. Barth, J.C. Björklund, M. Hoveskog, P.-O. Ulvenblad, and J. Ståhl. 2018. Barriers to business model innovation in the agri-food industry: A systematic literature review. *Outlook on Agriculture* 47, no. 4: 308–314.
Varman, R., and M. Chakrabarti. 2011. Notes from small industry clusters: Making sense of knowledge and barriers to innovation. *Ai & Society* 26: 393–415.
Vercher, N., G. Bosworth, and J. Esparcia. 2023. Developing a framework for radical and incremental social innovation in rural areas. *Journal of Rural Studies* 99: 233–242.