The influence of agricultural land structure on the identification of typical family dairy farms in the Republic of Croatia

Maja Petrač[1], Krunoslav Zmaić[1], David Kranjac[1], Lucija Pečurlić[1] and Jaka Žgajnar[2]

[1] Faculty of Agrobiotechnical Sciences Osijek, University of Osijek, Vladimira Preloga 1, Osijek, Croatia (mpetrac@fazos.hr)

[2] Biotechnical Faculty, University of Ljubljana, Jamnikarjeva 101, Ljubljana, Slovenia

Abstract

This paper presents the results of a cluster analysis conducted to identify typical dairy family farms in Croatia, based on four key variables related to land use. The objective is to determine the most representative farm types that can serve as a foundation for analysing the current state of the Croatian dairy sector. The analysis specifically investigates whether to allocate a greater proportion of land to pastures for grazing or to arable land for cultivating fodder crops. It also identifies the most commonly grown crops used for animal feed. Using these insights, we will develop farm models supported by a mathematical programming approach. The analysis is based on real-world data provided by the Croatian Agency for Agriculture and Food. It employs both hierarchical and non-hierarchical clustering methods using IBM SPSS Statistics software. The results offer a valuable basis for the development of farm-level models and support further research and evidence-based policy planning in the Croatian dairy sector.

INTRODUCTION

The long-term decline in the number of dairy farms, livestock, and overall milk production reflects an ongoing consolidation process typical of the agricultural sector (Gonzalez-Mejia et al., 2018), which is taking place across the entire European Union, including in Croatia. At the same time, the adoption of new technologies has contributed to an increase in milk yield per cow (Mijić et al., 2021). While small-scale farms are the most affected, often transitioning to arable farming, larger farms have recorded a modest increase in production (Očić et al., 2023). Given the unfavourable conditions and the limited effectiveness of current policy measures, a comprehensive analysis of the sector is required, alongside the development of more effective strategies for recovery. Specifically, there is a need for decision-support models that can offer deeper insight into the state of agricultural holdings and guide evidence-based policymaking (Ciaian et al., 2013). Strategic plans under the EU CAP increasingly incorporate simulation models at both the farm-level and aggregated scales (Lovec et al., 2020).

Before developing a representative model for Croatian dairy farms, it is essential to identify and classify typical farms based on shared characteristics and production capacities. Grouping dairy farms with similar profiles will enable a more accurate and relevant analytical approach (Chibanda et al., 2022). Poczta et al. (2020) identified five types of dairy farms in the EU using cluster analysis of FADN data. Croatian farms, according to average indicators, are classified into the first type, together with Slovenia, Austria, Poland, and Romania. These countries share a number of common structural and economic characteristics that clearly differentiate them from other members of the European Union. Cluster analysis is a widely used method for identifying groups of representative farms, or so-called typical farms (Pečnik et al., 2022). However, the outcome of such an analysis is not definitive, as it can vary depending on several factors within the analytical process, such as whether a hierarchical or non-hierarchical method is used, or which specific approach is applied within hierarchical clustering. Additionally, the choice of variables plays a crucial role in shaping the final grouping. Therefore, it is essential to carefully assess and justify the inclusion of each variable before including it in the analysis, as it directly influences how similarities between farms are measured.

In this paper, we present a set of variables related to the purpose of agricultural land use on dairy farms. Agricultural land is a key resource for dairy farms, serving not only as the primary source of feed for livestock but also as a foundation for economic profitability and environmental sustainability. The main objective is to analyse how agricultural land is utilised, with a specific focus on its primary function, by classifying dairy farms according to the dominant use of their land. This approach allows for a more detailed insight into the size, distribution, and purpose of agricultural land on farms specialized in milk production, factors that significantly influence the technology and efficiency of animal feed production.

The aim is to assess whether family dairy farms allocate more land to grazing or crop production, and to identify the main feed crops used. As feed systems impact both herd nutrition and manure management, land use strategies must support sustainability and address food security and environmental goals (N.P. Martin, 2017). The non-uniqueness of cluster analysis results stems from the choice of algorithm, distance metrics, and initial conditions can lead to different but equally valid solutions, The derived groupings will serve as a preliminary step in further research aimed at defining typical dairy farms. These will include not only land use characteristics but also production parameters and technology aspects. The final farm types will be validated and refined through expert workshops with agricultural consultants and experts, and further developed using the Slovenian SiTFarm model (Žgajnar et al., 2022). SiTFarm is a microsimulation tool based on mathematical programming that is an example of a bioeconomic farm model (BEFM).

MATERIALS AND METHODS

The empirical data used in this analysis were obtained from the Croatian Agency for Agriculture and Food (HAPIH). The data are collected from all agricultural holdings in Croatia that supply milk, with farmers reporting in a standardized format with relevant operational details. The initial database consisted of 4,198 registered dairy farms, recorded at the individual farm level. These records were supplemented with additional data from the Agency for the year 2022. During data preparation, efforts were made to resolve issues such as multiple identifiers, duplicate entries, and erroneous records (e.g., farms reporting zero cows). After thorough data cleaning and reconciliation, the final dataset comprised 3,393 dairy farms. Of these, 3,331 are categorized as family farms, while 62 are registered as legal entities. This paper focuses exclusively on the analysis of agricultural land use among family farms.

Statistical data processing was performed using the IBM SPSS Statistics V22.0 software package. Descriptive statistics for the variables used in cluster analysis are presented in Table 1. Unlike the previous cluster analysis on these data (Petrač et al, 2023) when the quantitative variables Number of cows (NOC), Annual delivery of milk (ADOM), Number of plant cultures (NOPC) and Area under culture (AUC) and Region (REG), (Petrač et al, 2024), the present study focuses on a different set of quantitative variables: AL - arable land for growing crops, PL - pasture land designated for cow grazing, ML - meadow land and OP - agricultural land for other purposes (e.g. vineyards, orchards, nurseries).

For each of these four variables, key descriptive statistics were calculated, including the mean, standard deviation (SD), minimum (Min), maximum (Max), and interquartile range (Q1–Q3), as summarized in Table 1.

Table 1. Descriptive statistics for family dairy farms

Variable Name	Variable	Mean	SD	Min	Max	Q1	Q2	Q3
ML	meadow land (ha)	3.11	5.31	0.00	131.32	0.49	1.79	3.90
AL	arable land (ha)	19.27	27.93	0.00	436.91	5.14	10.52	22.25
PL	Pasture land (ha)	0.67	5.01	0.00	163.80	0.00	0.00	0.00
OP	land with other purposes (ha)	0.63	5.60	0.00	169.94	0.00	0.00	0.09

The cluster analysis was first carried out on all the mentioned variables. Initially, a hierarchical (agglomerative) clustering approach was performed using Ward's method. This was followed by a non-hierarchical clustering approach, i.e., k-means algorithm, which was performed (Scitovski et al, 2022). The squared Euclidean distance was chosen as the distance measure. All algorithms were applied to standardized data. Among numerous solutions, one solution was selected as the most suitable for further analysis. Based on the visual interpretation of the dendrogram, the optimal number of clusters was identified and subsequently used as the input value for the k-means method.

RESULTS AND DISCUSSION

In 2022, family dairy farms delivered a total of 244,021,361.81 kg of milk, with an average delivery of 73,257.69 kg per farm, which accounts for 57.97% of the total national milk delivery in Croatia in that year. These farms cultivated a total of 78,885.98 ha, which represents 38% of the total agricultural land used by dairy farms in Croatia. The average land size per dairy farm was 23.68 ha, a significant increase from 2.7 ha in 2002 (Bosnić et al, 2003). One reason could be the trend of farm consolidation, i.e., many smaller farms have merged or been taken over by larger ones, which has resulted in an increase in the average size of farms and their arable areas.

According to the intended distribution of land, arable land dominates, occupying 64,189.63 ha (81.37%), followed by meadows (10,350.09 ha; 13.12%), continental grassland (2,246.30 ha; 2.85%), and a minor share of other land uses (2.66%) – including karst pastures (2.17%), orchards (0.30%), vineyards (0.13%), olive groves (0.05%), and mixed perennial plantations (0.01%) (Table 2). The majority of land is concentrated in Pannonian Croatia (according to Level 2 Statistical Regions (HR NUTS 2)), consistent with expectations based on its agricultural capacity.

Table 2. Distribution of land by purpose and NUTS regions of family dairy farms

	NOC	%	ML (ha)	AL (ha)	PL (ha)	OP (ha)	AUC (ha)	%
City of Zagreb	232	0%	118.51	263.73	0.47	1.85	384.56	0%
Pannonian Croatia	27,436	56%	4,940.71	39,895.42	1,846.72	481.30	47,164.14	60%
Northern Croatia	19,042	39%	4,369.45	21,265.06	396.67	144.00	26,175.19	33%
Adriatic Croatia	2,536	5%	921.42	2,765.42	2.44	1,472.81	5,162.09	7%
Total	**49.246**	**100%**	**10,350.09**	**64,189.63**	**2,246.30**	**2,099.96**	**78,885.98**	**100%**

Among the field crops, maize is the most widely grown crop, cultivated on 28,045.10 ha (44%) (Figure 1). Its dominance is not surprising as it is a key energy source in animal feed and is indispensable in silage production, which is particularly important for the nutrition of dairy cows during the winter. In addition to maize, significant crops include winter wheat (12%), grasses and forage (10%), alfalfa (9%) and winter barley (5%), all of which can serve as additional sources of animal feed and play a key role in providing nutrients needed for milk production. It is important to note that the upcoming weather conditions will make it difficult to

maintain cropping systems with limited diversity (i.e., monocultures or simple annual crop rotations in two phases). Therefore, grasslands could become an even more important way of diversifying agricultural systems (Sanderson et al., 2009).

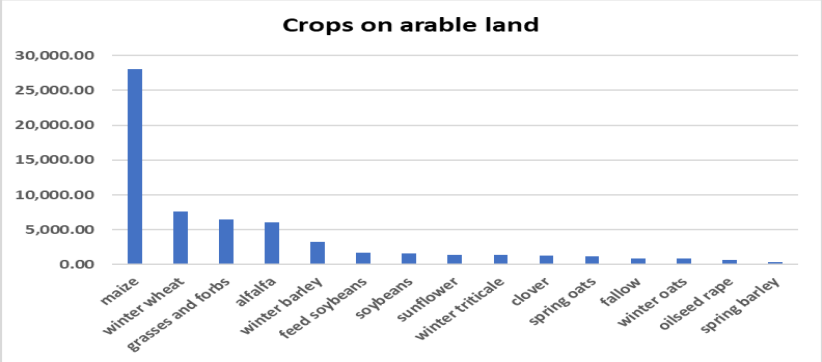

Figure 1. Crop distribution on arable land of family dairy farms

Further, we briefly present the results of the cluster analysis. First, we show the dendrogram to illustrate the hierarchical structure among the farms. Then, we present the results of the k-means clustering, which divides the family farms into 15 distinct clusters. For each cluster, we identify which farms belong to it, describe its key characteristics, and report how many farms it contains. We also highlight the average size of meadows, arable land, and other land types within each cluster. This allows us to see how land is represented across the different groups, which is the primary aim of this analysis.

The initial step in the cluster analysis involved determining the optimal number of clusters through the examination of a dendrogram (Figure 2), constructed using Ward's hierarchical method and the squared Euclidean distance as the dissimilarity measure.

The dendrogram shows how farms group based on similar characteristics. As the number of clusters decreases, the variation within each cluster increases. While the dendrogram provides guidance, the final choice of the number of clusters depends on expert judgment. Based on this assessment, the analysis determined that a 15-cluster solution most appropriately captures the diversity among the observed farms. The number of clusters was determined by analyzing changes in linkage heights within the dendrogram, generated using Ward's method, where pronounced discontinuities in the hierarchical structure indicated the optimal level for cutting the tree.

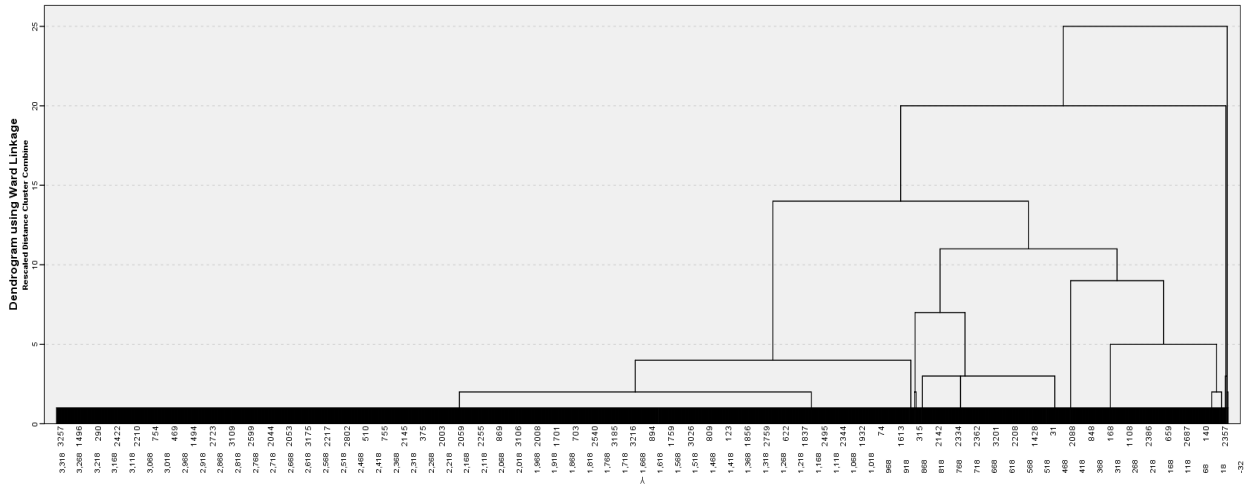

Figure 2. Dendrogram of family dairy farms according to variables (ML, AL, PL, OP)

Table 3. Cluster structure family dairy farms

Cluster	Number of farms	Average NOC	Average ML	Average AL	Average PL	Average OP	Average NOPC	Average AUC	Yield per cow
unit of measurement			ha	ha	ha	ha		ha	kg/cow
1	2,580	9.95	2.14	9.94	0.29	0.41	5.95	12.79	4,102.20
2	436	28.70	2.29	48.87	0.69	0.18	7.39	52.03	5,229.89
3	219	22.36	14.88	24.47	1.12	2.30	7.30	42.77	4,601.36
4	60	67.30	3.13	141.84	1.20	1.17	8.53	147.33	6,538.15
5	14	15.36	2.66	20.79	0.19	4.38	8.14	28.03	4,752.60
6	8	64.13	3.34	97.43	54.08	0.11	7.63	154.95	3,635.45
7	4	270.50	6.31	360.64	8.86	0.00	7.25	375.81	6,831.80

8	2	10.50	0.00	32.29	0.00	22.75	5.50	55.04	4,828.21
9	2	16.00	0.97	39.44	0.00	13.82	6.50	54.22	3,751.04
10	1	106.00	111.96	225.09	0.00	60.18	8.00	397.23	7,105.20
11	1	21.00	5.76	21.81	109.57	0.00	6.00	137.14	1,338.29
12	1	20.00	19.88	230.88	91.48	0.00	11.00	342.24	547.20
13	1	23.00	131.32	53.49	0.00	169.94	6.00	354.75	4,099.39
14	1	37.00	0.60	141.39	32.42	22.41	15.00	196.82	4,889.86
15	1	47.00	13.11	34.86	163.80	0.00	8.00	211.77	599.15
Grand Total	**3,331**	**14.78**	**3.11**	**19.27**	**0.67**	**0.63**	**6.29**	**23.68**	**4,329.80**

Legend: NOC-Number of cows, AL-arable land for growing crops, PL-pasture land designated for cow grazing, ML-meadow land, OP-agricultural land for other purposes, NOPC-Number of plant cultures, AUC-Area under culture

After deciding that there would be 15 clusters, using the k-means algorithm, the farms were distributed into 15 clusters. Table 3 shows the structure of all clusters after the implementation of k-means.

As can be seen from the table, 2,580 farms (77%) belong to cluster 1. This is the largest cluster by the number of farms, the average number of cows in this cluster (9.95) is less than the average number of cows in the Republic of Croatia (14.78), the average land area on the farm (12.79 ha) is less than the national average of 23.68 ha, the average area of meadows, arable land, pastures and other (2.14 ha, 9.94 ha, 0.29 ha and 0.41 ha) is also less than the national average (3.11 ha, 19.27 ha, 0.67 ha and 0.63 ha). This implies that most family dairy farms are located in cluster 1, which consists of very small farms with little arable land, meadows, pastures, and other.

Cluster 2 comprises 436 farms (13%) and bears some resemblance to cluster 1 in structure, however, farms in this group have much more agricultural land than farms in cluster 1. These are farms that have an average of 48.87 ha of arable land and 28.7 cows, and as such are larger than the average farm in the Republic of Croatia. Despite this, the relatively moderate herd size suggests that dairy farming may not be the primary production focus for these holdings.

Cluster 3 includes 219 farms (7%). In this cluster, too, farms have larger areas of agricultural land than the average in the Republic of Croatia. However, unlike cluster 2, the larger share of meadows is particularly notable here – the average area of meadows is 14.88 ha, which is many times higher than the national average of 3.11 ha. In addition, the average area of arable land is 24.47 ha. The average number of cows is not large, but it can still be concluded that these are farms that are primarily not engaged in dairy farming.

Cluster 4 stands out as the only group with average values exceeding national benchmarks across all key indicators. Consisting of 60 farms (2%), this cluster exhibits an average of 67.3 cows and 147.33 hectares of total agricultural land per farm—substantially higher than the national average of 23.68 hectares. Notably, these farms also demonstrate superior productivity, with an average milk yield of 6,538.15 kg per cow, which is higher than the Croatian average (4,329.8 kg). These attributes indicate a high level of specialization in dairy production and signal potential for further development and investment, consistent with findings by Žgajnar and Kavčič (2024) regarding high-performing Slovenian dairy farms.

From cluster 5 onwards, most clusters contain relatively few farms, often representing unique or outlier cases. Cluster 7 is interesting as it also includes four farms that are above-average farms in terms of both the average number of cows and the size of agricultural land as well as milk yield per cow, and cluster 10, which consists of one large farm.

The other clusters are isolated cases, consisting of 1 farm with a lot of arable land and few cows and a relatively low average production per dairy cow, which limits the scope for greater investments needed for the further growth and development of these farms. For this study, clusters comprising isolated cases were not subject to detailed analysis. These cases will be examined in the subsequent phase, following the development of the aforementioned model. Parzonko et al. (2024) state that in the EU, including Croatia, there is a great diversity of dairy farms in terms of scale and production technology.

CONCLUSION

This study highlights the diversity and structural complexity of Croatian family dairy farms, emphasizing agricultural land use as a key factor in their classification. Despite a continued national decline in the number of dairy farms and total milk production, the findings reveal that a significant number of family farms remain active, though often constrained by limited land resources and production capacity.

Cluster analysis identified 15 distinct groups, with the vast majority (77%) concentrated in a single cluster characterized by small herd sizes and minimal agricultural land. In contrast, only a few clusters represent larger, more productive farms with sustainable development potential. These results underscore the dual nature of the Croatian dairy sector: a predominant base of small-scale, low-yield farms alongside a smaller group of specialized, high-performing operations.

By categorizing farms according to land use patterns—particularly the distribution between arable land, pastures, and meadows—this research offers a robust basis for developing tailored, evidence-based policy measures. Furthermore, the classification framework serves as a preliminary step toward the development of representative farm models, which can inform strategic planning, guide investment decisions, and support long-term sustainability in the Croatian dairy sector.

REFERENCES

Bosnić, P., Mikulić, D., Vujčić, T., Marković, B. and Knjaz, V. (2003). Milk production in Croatia. Agriculturae Con-spectus Scientificus 68, 203-207.

Ciaian, P., Espinosa, M., Gomez Y. Paloma, S., Heckelei, T., Langrell, S., Louhichi, K., Sckokai, P., Thomas, A. and Vard, T. (2013). Farm level modelling of the CAP: a methodological overview. Publications Office of the European Union.

Chibanda, Craig, Katrin Agethen, Claus Deblitz, Yelto Zimmer, Mohamad. I. Almadani, Hildegard Garming, Christa Rohlmann, Johan Schütte, Petra Thobe, Mandes Verhaagh, and et al. (2020). The Typical Farm Approach and Its Application by the Agri Benchmark Network. Agriculture 10, no. 12: 646. https://doi.org/10.3390/agriculture10120646

Croatian Agency for Agriculture and Food. 2023. Annual report for 2022. Centre for Livestock Breeding, Cattle breeding

Croatian Bureau of Statistics. National classification of statistical regions 2021 (HR_NUTS 2021). National Gazette (in Croatian) (125/2019). Retrieved 2023-10-10.

Gonzalez-Mejia, A., Styles, D., Wilson, P. and Gibbons, J. (2018). Metrics and methods for characterizing dairy farm intensification using farm survey data. PLOS ONE, 13(5), e0195286. https://doi.org/10.1371/journal.pone.0195286

Hatfield, J. L., L. Wright-Morton and M. B. Hall (2017). Vulnerability of grain crops and croplands in the Midwest to climatic variability

and adaptation strategies. Clim. Change. https:// doi .org/ 10 .1007/ s10584 -017 -1997 -x.

Lovec, M., Šumrada, T. and Erjavec, E. (2020). New CAP Delivery Model, Old Issues. Intereconomics, 55(2), 112-119. https://doi.org/10.1007/s10272-020-0880-6

Malcolm, S., E. Marshall, M. Aillery, P. Heisey, M. Livingston and K. Day-Rubenstein (2012). Agricultural Adaptation to a Changing

Climate: Economic and Environmental Implications vary by U.S. Region. ERR-136. USDA-ERS, Washington, DC

Mijić, P. and Bobić, T. (2021). Milk production and challenges in transition from conventional to robotic milking in Croatia. Acta Scientiarum Polonorum Zootechnica, 20(3), 59-64. https://doi.org/10.21005/asp.2021.20.3.08

N.P. Martin, M.P. Russelle, J.M. Powell, C.J. Sniffen, S.I. Smith, J.M. Tricarico and R.J. Grant (2017). Invited review: Sustainable forage and grain crop production for the US dairy industry, Journal of Dairy Science, Volume 100, Issue 12, 2017, Pages 9479-9494, ISSN 0022-0302, https://doi.org/10.3168/jds.2017-13080

Očić, V., Šakić Bobić, B. and Grgić, Z. (2023). Economic analysis of specialized dairy farms in Croatia according to FADN. Mljekarstvo, 73(1), 50-58. https://doi.org/10.15567/mljekarstvo.2023.0106

Parzonko, A., Parzonko, A. J. and Bórawski, P. (2024). Economic Competitiveness of Dairy Farms from the Top Milk-Producing Countries in the EU: Assessment in 2014-2021. Agriculture, 14(1), 123. https://doi.org/10.3390/agriculture14010123

Pečnik, Ž. and Žgajnar, J. (2022). Resilience of dairy farms measured through production plan adjustments. Journal of Central European Agriculture, 23(1), 207-219. https://doi.org/10.5513/jcea01/23.1.3372

Petrač, M., Zmaić, K. and Žgajnar J. (2023). Typical family dairy farms in the republic of Croatia. 17th International Symposium on Operational Research in Slovenia, 31-34, Bled, Slovenia, September 20-22, 2023.

Petrač, M., Zmaić, K. and Žgajnar J. (2024). Typical family dairy farms in the republic of Croatia. 17th International Symposium on Operational Research in Slovenia, 31-34, Bled, Slovenia, September 20-22, 2023.

Poczta, W., Średzińska, J. and Chenczke, M. (2020). Economic Situation of Dairy Farms in Identified Clusters of European Union Countries. Agriculture, 10(4), 92. https://doi.org/10.3390/agriculture10040092

Sanderson, M. A., D. Wedin, and B. Tracey. (2009). Grassland: Definition, origins, extent, and future. Pages 57–74 in Grassland, Quietness and Strength for a New American Agriculture. W. F. Wedin and S. L. Fales, ed. American Society of Agronomy, Crop Science Society of America and Soil Science Society of America. Madison, WI.

Scitovski, R. and Sabo, K. (2020). Klaster analiza i prepoznavanje geometrijskih objekata. Sveučilište Josipa Jurja Strossmayera u Osijeku, Odjel za matematiku

Žgajnar, J., Kavčič, S., Tomšič, M., Zagorc, B., Brečko, J., Hiti Dvoršak, A., Moljk, B., Verbič, J., Bedrač, M., Kožar, Maja, Cunder, T. and Jerič, D. (2022). Razvoj modela za sistematično spremljanje ekonomskega položaja in analizo vpliva kmetijske politike na ravni tipičnih kmetijskih gospodarstev: zaključno poročilo. Ljubljana: Biotehniška fakulteta, feb. 2022. 437 str., ilustr. [COBISS.SI-ID 97899267]

Žgajnar, J. and Kavčič, S. (2024). Understanding the Dairy Sector in Slovenia: A Modeling Approach for Policy Evaluation and Decision Support. Sustainability, 16(14), 6009. https://doi.org/10.3390/su16146009