The influence of the region on the determination of typical family dairy farms in the Republic of Croatia

Petrač, M.; Zmaić, Krunoslav; Kranjac, David; Mijić, Pero; Žgajnar, J.

Sveučilište Josipa Jurja Strossmayera u Osijeku

Fakultet agrobiotehničkih znanosti Osijek

Repository / Repozitorij:

Repository of the Faculty of Agrobiotechnical Sciences Osijek - Repository of the Faculty of Agrobiotechnical Sciences Osijek

SVEUČILIŠTE JOSIPA JURJA STROSSMAYERA U OSIJEKU

dabar

DIGITALNI AKADEMSKI ARHIVI I REPOZITORIJI

The influence of the region on the determination of typical family dairy farms in the Republic of Croatia

Maja Petrač[1], Krunoslav Zmaić[1], David Kranjac[1], Pero Mijić[1], Jaka Žgajnar[2]

[1]*Josip Juraj Strossmayer University of Osijek, Faculty of Agrobiotechnical Sciences 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 cluster analysis to determine typical farms, in which the variable region classified by NUTS2 was added in order to analyze the influence of that variable. The aim is to find the most representative typical dairy family farms that could be used to analyse the situation in the dairy sector in Croatia. On the basis of these information, we will further create farm models supported by the mathematical programming approach. Cluster analysis was performed on real data obtained from the Croatian Agency for Agriculture and Food. Hierarchical clustering and non-hierarchical clustering were performed using IBM SPSS Statistics.

Keywords: cluster analysis, typical farms, dairy sector, NUTS2 regions, farm model, Croatia

Introduction

The dairy sector in the Republic of Croatia has been characterized by negative trends for years, as shown by the decline in the number of farms, livestock, and the amount of milk production, while a positive effect is expected in the increase of milk yield per cow due to the introduction of new technologies (Mijić et al., 2021). During 2022, the decreasing trend continues. In the Republic of Croatia 405 425 t of milk were purchased, i.e. by 5,4% less than the previous year, with a decrease in the number of suppliers by 13,7%. A positive trend is the increase in delivered quantities per supplier by 12,3% (Croatian Agency for Agriculture and Food, 2023). Small farms with few animals are either closed or transferred to arable production (Mijić et al., 2021). The Croatian dairy sector is not expected to recover soon, and the simulation results indicate a further decline in the number of dairy cows and the amount of milk produced (Kranjac, 2020). The situation has been further aggravated in the last few years due to the consequences of the corona crisis and the war in Ukraine. The measures implemented obviously do not lead to improvement of the situation.

Before the actual creation and application of the model at the farm level, it is necessary to determine typical dairy farms in Croatia, i.e. farms should be grouped according to common characteristics (Chibanda et al., 2020). For this purpose cluster analysis could be used to obtain these groups of representative farms which are called typical farms (Pečnik et al., 2022). Cluster analysis solutions are not unique and depend on the application of different elements of the analytical procedure (e.g. hierarchical or non-hierarchical method, different methods of the hierarchical method). The final solution also depends on the variables that are used as a basis for measuring similarity of the group, so one should be careful about the impact of each decision when choosing variables.

In this paper we present one of the possible combination of variables of cluster analysis that has been conducted. Since the results of the cluster analysis are not unique, the cluster analysis in our holistic analysis of dairy sector will serve as the starting point for further analysis and defining final typical farms (production parameters, technologies etc.), which will further be modeled with the farm model. Typical farms will be defined in more detail at workshops with consultants and experts in the field and will be further adjusted and upgraded with the Slovenian farm model - SiTFarm (Žgajnar et al., 2022). For the workshop with consultants, several variants of cluster analysis will be created by changing the specified parameters.

This paper describes the procedure of cluster analysis for the case when, in addition to the previous 4 variables already analysed (Petrač et al., 2023) (Number of cows (NOC), Annual delivery of milk (ADOM), Number of plant cultures (NOPC) and Area under culture (AUC)), the variable Region (REG) was also added. This provides additional insight into where individual farms are located, and the technology of animal husbandry and technology of fodder production also depends on this.

Materials and methods

Cluster analysis was performed on real data obtained from the Croatian Agency for Agriculture and Food. The database consisted of 4198 dairy farms that supply milk in the Republic of Croatia. After arranging the obtained database (connecting data from different farm databases, removing duplicate and inactive farms, etc.) there were 3398 farms left for the analysis. Since in this analysis we focus on family farms, 67 of the farms have been excluded from the sample, since they have the status of a legal entity and form a special category. A separate cluster analysis will be done on them. Therefore, the final number of farms analysed is 3331. These are family farms that delivered milk and as such were included in the final register (Table 1).

IBM SPSS, Statistics V22.0 software package was used for statistical data processing and analysis. The variables and their descriptive statistics are presented in Table 1. Unlike the previous cluster analysis on these data (Petrač et al,, 2023) when the quantitative variables NOC, ADOM, NOPC and AUC were used, here the qualitative variable Region (REG) was added. To define the mentioned variable, the national classification of statistical regions from 2021 was used - Level 2 Statistical Regions (HR NUTS 2), in which the Republic of Croatia is divided into 4 regions (Croatian Bureau of Statistics, 2021): Pannonian Croatia, Adriatic Croatia, City of Zagreb and Northern Croatia (Figure 1).

Figure 1. Level 2 Statistical Regions (HR NUTS 2)

Quantitative variables and their descriptive statistics are presented in Table 1. Several numerical characteristics of the variables (mean, standard deviation (SD), minumum (min) and maximum (max)) were determined for these 4 quantitative variables.

Table 1. Descriptive statistics for quantitative variables - 3.331 dairy farms

Variable Name	Variable	Mean	SD	Min	Max
NOC	Number of cows	14.78	20.47	1	456
ADOM	Annual delivery of milk (kg)	73 257	146 015	21	2 799 071
NOPC	Number of plant cultures	6.29	2.27	1	18
AUC	Area under culture (ha)	23.68	31.68	0.15	469.05

Source: Own calculations

The relative frequency was determined for the qualitative variable REG (see Figure 2). The most dairy farms are located in the Pannonian region (47%) and Northern Croatia (45%). Both regions are located in the continental part of the country and include eight and five counties of Croatia, respectively. The Adriatic region forms the coastal part of the country, including seven counties and 7% of farms are located in it. As expected, the least farms are located in the city of Zagreb, with about 1%.
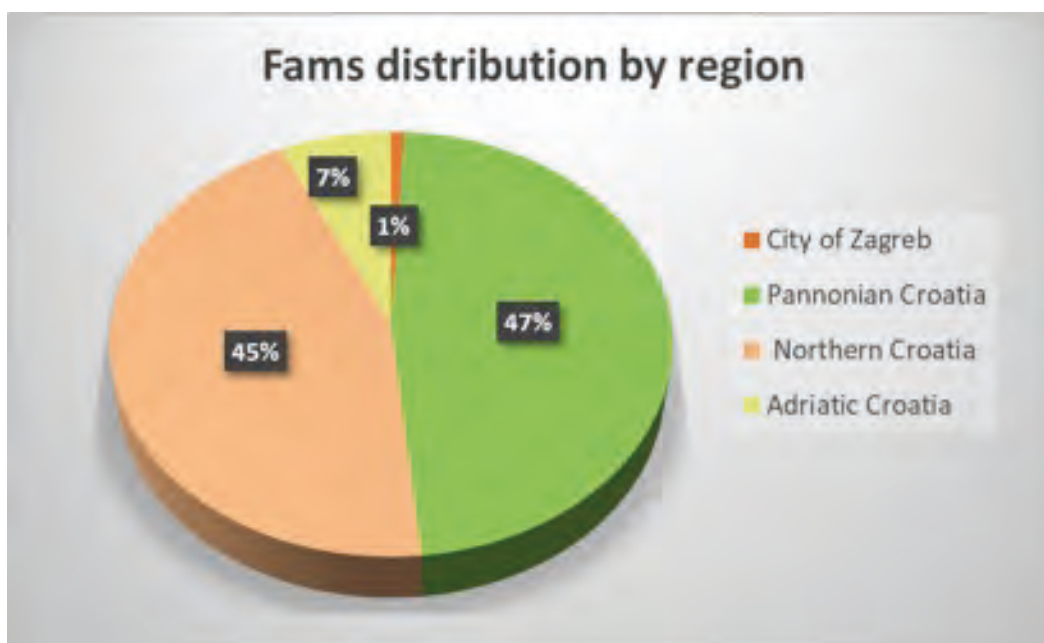

Figure 2. Descriptive statistics for qualitative variable (relative frequency)

The cluster analysis was first carried out in relation to all the mentioned variables. First, hierarchical (agglomerative) clustering was performed using Ward's method. The results of the agglomerative method can be graphically displayed in the form of a two-dimensional hierarchical diagram, known as a dendrogram. Then non-hierarchical clustering, i.e., k-means algorithm, was performed (Scitovski et al., 2020). The squared Euclidean distance was chosen as the distance measure. All algorithms were applied to standardized data. Among numerous solutions, finally one solution was chosen as the final one. Cluster analysis solutions are not unique and depend on the application of different elements of the analytical procedure (e.g. hierarchical or non-hierarchical method, different algorithms of the same method). One should be careful about the impact of each dec ision when choosing variables becouse the solution also depends on the variables that were used as a basis for measuring similarity. The final decision remains for users (the researchers and expert) to estimate. The experts will be the ones who will choose the cluster that best describes the real situation in the sector.

Results and discussion

Further we briefly present the main results. In the first part, we show the dendrogram, and further the results of the k-means algorithm will be shown, where we can see how the farms are distributed in 15 clusters, i.e. which farms belong to which cluster and what are its characteristics, how many farms are in which cluster, what is the average number of cows in each cluster as was the goal of this work, we see the representation of regions in each cluster.

The first step in this analysis was to determine the optimal number of clusters using the dendrogram (see Graph 1) using Ward's method, based on the squared Euclidean distance. The structure of the data was analyzed, i.e. different groups of farms were found, which share common characteristics. How heterogeneity within clusters increases, so the number of clusters decreases. The dendrogram helps in deciding on the optimal number of clusters, but the final decision remains for the researchers and expert to estimate. As apparent from the dendrogram, analysis suggests 15 clusters.
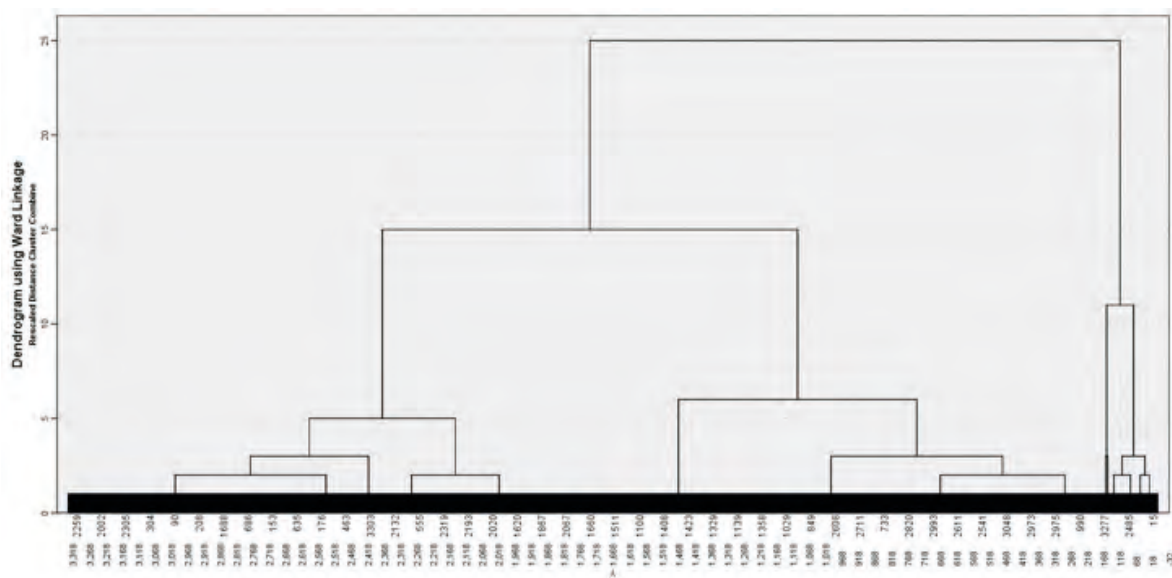
Graph 1. Dendrogram

After it was decided that there would be 15 clusters, using the k-means algorithm, the farms were distributed into 15 clusters. Table 2a and 2b show the structure of all clusters after the implementation of k-means.

As apparent from Table 2a, 836 farms (25%) belong to cluster 1. The average number of cows in this cluster (8.01) is smaller than the average number of cows in the Republic of Croatia (14.78), the average annual milk delivery (33870.71 kg) is lower than the national average of 73257.69 kg, the average area of land per farm (10.48 ha) is also smaller than the national average. This cluster consists mainly of farms from Northern Croatia (90%), and a small part of farms is from Adriatic region (10%). This implies that cluster 1 consists mainly of very small farms from Northern Croatia.

It can be seen that cluster 2 consists of 653 farms (20%) and is also relatively similar to cluster 1, but farms in cluster 2 have much more agricultural land as well as plant cultures than farms in cluster 1. Namely, there are many farms with a few cows and a lot of land, and these are not the farms whose primary activity is milk production. It can also be concluded that in this cluster the majority of farms are from Northern Croatia. However with a slightly smaller share of farms from Northern Croatia, and a larger share from Adriatic Croatia.

Cluster 3 consists of 621 farms (19%). The average number of cows in this cluster (9.84) is also smaller than the average number of cows in the Republic of Croatia (14.78), the average annual milk delivery (38071.60 kg) is lower than the national average of 73275.69 kg, the average area under culture (12.76 ha) is also smaller than the national average. Almost all of the farms in this cluster are from Pannonian region, and a very small part of farms are from City of Zagreb. This implies that cluster 3 consists of very small farms in Pannonian Croatia.

Again, cluster 4 is similar to cluster 3, with the only difference being that the farms in cluster 4 have much more

land and plant cultures. Namely, there are many farms in Pannonian region with a few cows and a lot of land, and these, obviously, are not the farms that are primarily active in milk production and dairy is not the only activity on these farms. From the fifth cluster onwards, there are smaller groups that represent 18% of the total number of farms. However, they are extremely important from the point of view of milk production. Even 57% of the total production of dairy cows takes place on these farms (Table 2a). With the exception of the fifth and eighth clusters, milk production per cow is significantly higher on all farms than in the first four clusters. The two largest farms are in special clusters (cluster 14 and 15). They are both located in Pannonian Croatia. The 6 largest Croatian farms (cluster 12 to 15) whose production is 5% in the total milk production in Croatia are also located in the Pannonian region (Table 2b).

Table 2a. Cluster structure

Cluster	Number of farms	Average NOC	Average NOPC	Average AUC	Yield per cow
1	836	8.01	4.75	10.48	4,186.47
2	653	9.27	8.22	15.12	4,103.21
3	621	9.84	4.01	12.76	3,917.06
4	611	11.08	7.02	21.84	4,055.33
5	221	18.75	10.11	48.25	4,529.97
6	213	32.41	6.15	39.43	5,600.62
7	93	56.49	7.01	67.82	6,621.60
8	35	33.60	8.83	175.84	4,423.23
9	32	88.66	7.44	119.00	7,797.23
10	5	166.80	5.80	98.52	7,747.71
11	5	204.00	8.20	226.84	6,950.93
12	2	230.50	4.50	204.51	9,123.60
13	2	107.50	7.50	388.93	6,309.97
14	1	317.00	8.00	469.05	8,829.88
15	1	456.00	6.00	370.27	5,614.38
Croatia	3331	14.78	6.29	23.68	4,329.80

Legend: NOC - Number of cows, NOPC - Number of plant cultures
AUC - Area under culture (ha)

Table 2a also contains data on milk yield per cow for each cluster (**Yield per cow** column). This data was calculated from the obtained data base and was not included as a variable in the cluster analysis.

Table 2b. Cluster structure by regions

Cluster	Number of farms	Number of farms (%)	ADOM (%)	City of Zagreb	Pannonian Croatia	Northern Croatia	Adriatic Croatia
1	836	25%	12%	0%	0%	90%	10%
2	653	20%	10%	0%	0%	80%	20%
3	621	19%	10%	1%	99%	0%	0%
4	611	18%	11%	3%	97%	0%	0%
5	221	7%	7%	0%	83%	17%	0%
6	213	6%	15%	0%	37%	60%	3%
7	93	3%	14%	0%	62%	33%	4%
8	35	1%	2%	0%	74%	14%	11%
9	32	1%	9%	0%	66%	28%	6%
10	5	0%	3%	0%	60%	40%	0%
11	5	0%	3%	0%	60%	40%	0%
12	2	0%	2%	0%	100%	0%	0%
13	2	0%	1%	0%	100%	0%	0%
14	1	0%	1%	0%	100%	0%	0%
15	1	0%	1%	0%	100%	0%	0%
Grand Total	3331	100%	100%				

Legend: ADOM - Annual delivery of milk

Conclusion

By applying cluster analysis to the data of dairy farms in Croatia, many different solutions could be obtained. By introducing the variable Region (REG), new and different solutions are obtained. As expected, the variable Region affects the new division of farms, and 15 clusters are obtained. Most of the farms are mainly distributed in the Northern and Pannonian regions. These are mostly small farms with low milk production. Since the farms of the Northern Region are mostly in hilly areas, they are limited in arable land availability. Unlike the Northern Region, the Pannonian region has a lot of arable land, so the farms in that area have more arable land on which to produce their own fodder for livestock and thereby reducing production costs. Therefore, it is not surprising that the largest farms are located in the same region, which are also the main carriers of the total milk production. Small farms are crucial for rural areas, especially from a socioeconomic aspect. On the other hand, they are not a key factor that contributes to the increase in production volume, food security, or the increase in market self-sufficiency. Typical farms have different levels of economic efficiency so the same measure is not equally effective for every typical farm. A certain policy measure that would increase the profitability of one typical farm, does not mean that it would increase the profitability of another type of farm. Such results could be useful to policy makers to get information on which typical dairy farm needs which type of measure and how much they can adapt to the given situation. Such more effective policy measures should encourage more economical farming.

References

Chibanda C., Agethen K., Deblitz C., Zimmer Y., Almadani M. I., Garming H., ... Lasner T. (2020). The Typical Farm Approach and Its Application by the Agri Benchmark Network. Agriculture. 10: 646.

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.

Kranjac D., Zmaić K., Salputra G., Salamon P., Erjavec E. (2020). Production and trade impacts of CAP past 2021 reform scenarios on main Croatian crop and livestock sectors. German Journal of Agricultural Economics.

Mijić P., and Bobić T. (2021). Milk production and challenges in transition from conventional to robotic milking in Croatia. Acta Sci. Pol. Zootechnica. 20: 59–64.

Pečnik Ž., and Žgajnar J. (2022). Resilience of dairy farms measured through production plan adjustments = Odpornost kmetij s prirejo mleka z različnimi prilagoditvami proizvodnega načrta. Journal of Central European Agriculture. 23: 207-219.

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.

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., 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]