

**KLĀSTERANALĪZES LIETOŠANA  
PAGASTU GRUPĒŠANAI  
CLUSTER ANALYSIS FOR THE  
CLASSIFICATION OF RURAL COMMUNITIES**

**A. Melluma**

LLU Videssaimniecības katedra  
Department of Environmental engineering and management, LUA

**L. Ramute, P. Rivža**

LLU Informātikas institūts  
LUA Institute of Informatics

**Abstract.** The article presents results of grouping Latvia's pagasts by cluster analysis producing 4 singular pagast groups. As variables 22 pagast descriptiv factors were used in different combinations.

**Key words:** rural communities, classification, cluster analysis.

## **1. Ievads**

Zinātnē un praksē pētnieki bieži sastopas ar sekojošu problēmu: kā labāk atrisināt objektu (piem., aptauju rezultātu, firmu vai dzīvnieku) sadalīšanu viendabīgās grupās. Šī uzdevuma atrisināšanai var lietot klāsteru analīzi.

Par klāsteru analīzi sauc tādu metožu kopumu, kuru galvenais, primārais mērķis ir, balstoties uz noteiktu raksturojošu parametru salīdzinājumu, sagrupēt pētāmos objektus vai individuus apakškopās jeb klāsteros. Jāatzīmē, ka katrs objekts tiek iekļauts vienā un tikai vienā klāsterā. Klāsteru analīze klasificē objektus tā, ka viena klāstera objekti pēc raksturojošiem parametriem ir līdzīgi, bet tai pašā laikā, objekti, kas pieder dažādiem klāsteriem, ir atšķirīgi. Var teikt, ka klāsteranalīzes uzdevums ir grupēšana, kas apmierinātu kādu optimalitātes kritēriju.

## **2. Materiāls un metodes**

Analizējot un plānojot Latvijas pagastu vai atsevišķu reģionu darbību, ir nepieciešams veikt visu pagastu vai atsevišķa reģiona pagastu sagrupēšanu viendabīgās grupās pēc vairākiem faktoriem, tādējādi pagasti kļūtu ērtāk salīdzināmi. To var izdarīt ar klāsteranalīzes metodēm. Jāatzīmē, ka Latvijas pagastu grupēšanai klāsteranalīze tiek izmantota pirmo reizi.

Izpētes objekts ir matrica, kas sastāv no 508 Latvijas pagastiem un 22 faktoriem, kas izvēlēti kā šo pagastu raksturotāji. Tie ir:

- kopplatība;
- iedzīvotāju skaits 1995. gada sākumā;
- X1 - iedzīvotāju skaits uz 100 ha lauksaimniecībā izmantojamās zemes;
- X2 - lauku iedzīvotāji darbības vecumā, procentos no kopskaita;
- X3 - iedzīvotāju blīvums, iedzīvotāji uz 1 km<sup>2</sup>;

- X4 - iedzīvotāju skaits ar pamatskolas vai zemāku izglītību, un analfabēti uz 1000 iedzīvotājiem;
- X5 - pensionāru īpatsvars iedzīvotāju skaitā, %;
- X6 - bezdarbnieku īpatsvars darbspējas vecumā esošo iedzīvotāju skaitā, %;
- X7 - izlīdzinātie ieņēmumi uz 1 iedzīvotāju, Ls/ 1 iedz.;
- X8 - izlīdzinātie izdevumi uz 1 iedzīvotāju, Ls/ 1 iedz.;
- X10 - ieņēmumu prognoze no iedzīvotājiem, Ls/ 1 iedz.;
- X11 - ienākuma nodoklis uz 1 iedzīvotāju, Ls/ 1 iedz.;
- X12 - vietējā budžeta izdevumi uz 1 iedzīvotāju, Ls/ 1 iedz.;
- X13 - vietējā budžeta izdevumi uz 1 darbspējīgo, Ls/ 1 darbsp.;
- X14 - dotācijas apjoms uz 1 iedzīvotāju, Ls/ 1 iedz.;
- X15 - lauksaimniecībā izmantojamās zemes (LIZ) īpatsvars pagasta teritorijā, %;
- X16 - LIZ nekoriģētais novērtējums, Ls/ha;
- X17 - LIZ koriģētais novērtējums, Ls/ha;
- X18 - transporta līdzekļu skaits uz 1000 iedzīvotājiem;
- X19 - pagasta tips pēc attāluma līdz rajona centram;
- X20 - pagasta tips pēc Latvijas robežas tuvuma;
- X21 - pagasta tips pēc atrašanās transporta koridora tuvumā.

Klāsterizācijas procesam ir divi algoritmi: hierarhiskais un nehierarhiskais, katram no tiem ir savas darba metodes. Hierarhiskajam algoritmam tiek izmantotas uzkrāšanas un dalīšanas metodes, nehierarhiskajam - secīgo klāsteru (Sequential Threshold), paralēlo klāsteru (Parallel Threshold) un optimizācijas (Optimization) metodes.

Klāsterizācijā varam izmantot vai nu vienu no iepriekš norādītajiem algoritmiem, vai abus. Latvijas pagastu grupēšanai izmantota nehierarhiskā algoritma paralēlo klāsteru (Parallel Threshold) metode, tā atļauj lietotājam, pamatojoties uz praktiskajiem un teorētiskajiem rādītājiem, norādīt klāsteru skaitu. Mūsu pētījumā izvēlētas 4 grupas jeb klāsteri.

Latvijas pagastu grupēšana tika veikta ar programpaketes SPSS for MS Windows (6.0 versijas) metodi QUICK CLUSTER. Rakstā aplūkoti sekojoši grupējumi, kuri atšķiras ar izmantotajiem pagastus raksturojošajiem faktoriem:

1. grupējumam izmantoti visi 22 pagastus raksturojošie faktori;
2. grupējumam izmantoti 9 no 22 faktoriem - X2, X3, X4, X5, X6, X7, X8, X15, X17;
3. grupējumam izmantoti 3 faktori - X3, X10, X17;
4. grupējumam izmantoti sekojoši faktori:
  - kopplatība;
  - iedzīvotāju skaits 1995. gada sākumā;
  - X3.
5. grupējumam izmantoti 4 faktori - X11, X17, X2, X6;
6. grupējumam lietoti:
  - kopplatība;
  - iedzīvotāju skaits 1995. gada sākumā;
  - X3;
  - X15.
7. grupējumam izmantoti 20 faktori - X1, X2, X3, X4, X5, X6, X7, X8, X10, X11, X12, X13, X14, X15, X16, X17, X18, X19, X20, X21.
8. grupējumam izmantoti faktori, kurus mēs lietojam arī faktoru analizē - X2, X3, X4, X5, X6, X11, X12, X14, X15, X17, X18, X19.

### 3. Rezultāti

Rezultāti, kas iegūti pēc veiktajiem 8 grupējumiem, apkopoti 1. tabulā. 4. un 6. grupējumā iegūti identiski pagastu sagrupējumi. Tas tādēļ, ka 4. un 6. grupējumā tika izmantoti tādi raksturojošie

1.tabula/Table 1

**Latvijas pagastu sadalījums klāsteros**  
**Classification of Latvian rural communities**

Grupējumi Groupings	Klāsteri Clusters							
	1.		2.		3.		4.	
	Frekvence Frequency							
	absolūtā	relatīvā %	absolūtā	relatīvā %	absolūtā	relatīvā %	absolūtā	relatīvā %
1.	23	4.5	318	62.6	166	32.7	1	0.2
2.	491	96.7	15	3.0	1	0.2	1	0.2
3.	293	57.7	3	0.6	168	33.1	44	8.7
4.	23	4.5	318	62.6	166	32.7	1	0.2
5.	48	9.4	170	33.5	1	0.2	289	56.9
6.	23	4.5	318	62.6	166	32.7	1	0.2
7.	1	0.2	489	96.3	2	0.4	16	3.1
8.	72	14.2	1	0.2	268	52.8	167	32.9

faktori, kuri būtiski ietekmē klāsterizācijas iznākumu, un tie ir:

- kopplatība;
- iedzīvotāju skaits 1995. gada sākumā;
- X3 - iedzīvotāju blīvums, iedzīvotāji uz 1 km<sup>2</sup>;
- X15 - lauksaimniecībā izmantojamās zemes (LIZ) īpatsvars pagasta teritorijā, %;

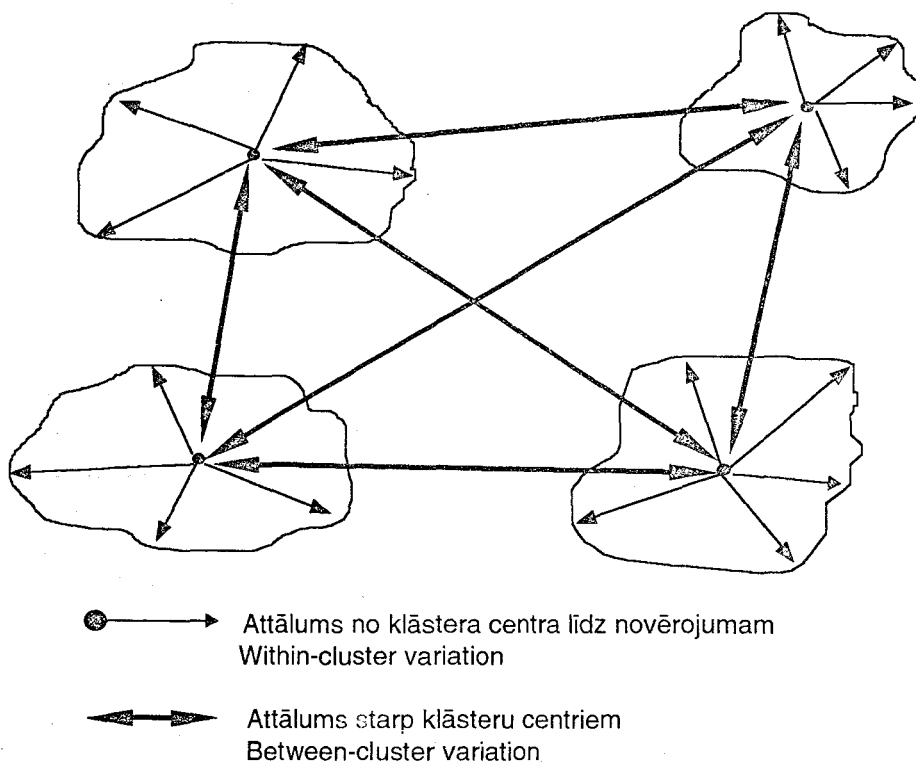
Redzam, ka starp šiem faktoriem nav neviena ekonomiskas dabas rādītāja; šajos grupējumos tie praktiski maz ietekmē pagastu klāsteru veidošanu. Ar to arī tiek apstiprināts, ka Latvijas pagastos, piemēram, ir zema ieņēmumu prognoze no iedzīvotājiem, neliels ienākumu nodokļu apjoms, nelieli vietējā budžeta izdevumi uz 1 iedzīvotāju.

Iepazīstoties ar grupējumu rezultātiem, redzam, ka kāds no pagastiem, krasi atšķiroties pēc kādas no pazīmēm, tiek atdalīts atsevišķā grupā. 1., arī 4. un 6., grupējumā tas ir Talsu rajona Dundagas pagasts (4. klāsterā), kurš atšķiras no citiem pagastiem ar savu kopplatības lielumu. 2. grupējumā kā atšķirīgi no pārējiem tika minēti Rīgas rajona Garkalnes pagasts (3. klāsterā) un Valmieras rajona Zilākalna pagasts (4. klāsterā). 5. grupējumā - Rīgas rajona Salaspils (3. klāsterā), arī 8. grupējumā - Rīgas rajona Salaspils (2. klāsterā). Varam norādīt, ka 3. un 5. grupējumu rezultāti atbilst pagastu vienmērīgākam sadalījumam klāsteros.

Jāpiezīmē, ka datorprogramma SPSS, bez objektu sadalīšanas pa klāsteriem, vēl dod sekojošus rādītājus: klāsteru centrus, attālumus no katra novērojuma līdz klāsteru centram, attālumus starp klāsteriem (skat. 1. att.), kā arī dispersiju analīzes rezultātus, pēc kuriem var analizēt raksturojošo faktoru svarīgumu klāsterizācijai.

#### 4. Slēdziens

Klāsteru analīze ieteicama Latvijas pagastu grupēšanai. Ļoti rūpīgi jāizvēlas klāsterizācijas mainīgie, lai tie visi būtu statistiski nozīmīgi. Ja klāsteru skaitu nevar noteikt pirms klāsterizācijas, tad jāizmanto hierarhiskā klāsterizācijas metode. Strādājot ar nehierarhisko klāsterizācijas metodi, pagastu grupu skaitu var izvēlēties pēc citiem apsvērumiem.



1. att. Shematisks klāstera attēlojums.

Fig.1. Cluster diagram showing between- and within-cluster variation.

## Literatūra

Multivariate data analysis: with readings. (1995). Joseph F. Hair, Rolph E. Anderson, Ronald L. Tatham, William C. Black. PRENTICE HALL, 4th edition. 420-483.