

Мкртчян О.С., Шубер П. М. Регіоналізація річного режиму атмосферних опадів західної частини України методом кластерного аналізу // Фізична географія і геоморфологія. – 2008. – Вип.54. – С.187-194.

Мкртчян О.С., Шубер П.М.

Львівський національний університет імені Івана Франка

РЕГІОНАЛІЗАЦІЯ РІЧНОГО РЕЖИМУ АТМОСФЕРНИХ ОПАДІВ ЗАХІДНОЇ ЧАСТИНИ УКРАЇНИ МЕТОДОМ КЛАСТЕРНОГО АНАЛІЗУ

Постановка проблеми. Клімат є однією з найважливіших фізико-географічних характеристик будь-якої території, що безпосередньо впливає на ґрунтовий покрив, рослинний і тваринний світ, гідрологічні умови, умови життєдіяльності та господарської діяльності людини. Важливою характеристикою клімату є режим атмосферних опадів, який визначає умови річкового стоку (у тому числі – формування паводків на річках), забезпеченість вологою ґрунтів, рослин тощо.

Кліматичне районування є поділом території на регіони, однорідні щодо основних кліматичних показників. Першим кроком до загального кліматичного районування має стати аналіз просторового розподілу основних складових кліматичного режиму, однією з яких є режим атмосферних опадів. Одним з шляхів такого аналізу є регіоналізація – виділення територіальних одиниць, внутрішньо однорідних за цим режимом.

Просторовий розподіл кількості атмосферних опадів суттєво різниться в залежності від сезону року та переважаючого режиму атмосферної циркуляції. Режим опадів впродовж конкретного року визначається їхньою загальною кількістю та розподілом за місяцями року.

Аналіз стану проблеми та методика досліджень. Для виділення просторових одиниць, однорідних у відношенні просторового розподілу атмосферних опадів, доцільно застосовувати метод кластерного аналізу. Метою кластерного аналізу є виявлення характерних структур у багатовимірних даних, а саме – поділ сукупності об'єктів, що описуються багатьма кількісними змінними, на групи, однорідні щодо розподілу цих змінних. В нашому випадку такими об'єктами є окремі метеостанції, а змінними – місячні суми опадів за певні роки.

Існує два різновиди методів кластерного аналізу: дедуктивні та ієрархічні. В першому випадку на початку усі об'єкти приписуються до наперед визначеної кількості кластерів випадковим способом, після чого об'єкти переміщуються між кластерами (з одночасним переміщенням центроїдів кластерів) доти, доки не буде досягнуто найбільшої гомогенності кластерів, тобто найбільшої подібності між об'єктами всередині кластерів та найбільшої відмінності між різними кластерами. Найпоширенішим з таких методів є метод к-середніх.

При використанні ієрархічних методів, натомість, на початку кожному об'єкту відповідає свій окремий кластер, після чого об'єкти (кластери), найбільш подібні між собою, об'єднуються у більші кластери, утворюючи, таким чином, ієрархічну структуру.

Для визначення порядку об'єднання об'єктів у кластери найчастіше обраховують Евклідову відстань між об'єктами за формулою $D_{ij} = \sqrt{\sum_n A_{ij}^2}$, де A – відмінність у значенні певного атрибуту об'єктів i та j , n – кількість атрибутів. Після утворення кластерів об'єднання останніх у більші кластери доцільно проводити, використовуючи метод Уорда [8]. Цей метод полягає у мінімізації суми квадратів відстаней між об'єктами у новоутворених кластерах. Таким чином виділяються найбільш внутрішньо однорідні кластери. Якщо дедуктивні методи (зокрема, метод k -середніх) виділяють наперед задану кількість кластерів, то ієрархічний кластерний аналіз послідовно зменшує цю кількість, об'єднуючи "вужчі" кластери у "ширші": таким способом, на перших етапах аналізу отримуємо велику кількість внутрішньо однорідних кластерів, які на наступних етапах об'єднуються у меншу кількість більш гетерогенних кластерів; цей процес продовжується, доки усі об'єкти не будуть поєднані у один спільний кластер.

Результати ієрархічного кластерного аналізу найкраще зобразити у формі дендрограми, яка зображує цю послідовність кластеризації об'єктів (див. рис. 1 та 3 нижче). Зліва кожному об'єкту відповідає окрема "гілка" дендрограми; об'єднання об'єктів у кластери зображається вузлом (злиттям гілок), причому відстань по горизонталі від початкової вертикальної осі до кожного вузла пропорційна відмінностям між кластерами, що зливаються (i , відповідно, внутрішній гетерогенності кластеру, що утворився). Таким чином, чим більш подібними між собою є об'єкти або кластери, тим лівіше на дендрограмі розташований вузол їх злиття. При використанні цього методу кількість кластерів, які отримують "на виході", як правило визначається суб'єктивно шляхом візуального аналізу дендрограми: чим довшою є гілка, яка відповідає певному кластеру, тим більш чітко вираженим є цей кластер.

Метод кластерного аналізу успішно використовувався з метою регіоналізації (виділення кліматичних районів з однорідним режимом атмосферних опадів) дослідниками ряду країн: США [4], Швейцарії [2], Ірану [3], Китаю тощо.

Заслугове на увагу дослідження іспанських геофізиків [7], в ході якого методом кластерного аналізу даних 32 метеостанцій територію країни було поділено на декілька регіонів, відмінних за режимом опадів. Було, зокрема, виявлено, що кількість, розміри та межі регіонів суттєво відрізняються за сезонами року, що пояснюється відмінностями у генезисі опадів. Результати, отримані з використанням кластерного аналізу, виявились подібними до результатів, отриманих з використанням методу головних компонент – ще однієї поширеної статистичної методики структурного аналізу кількісних даних.

Метою даного дослідження є регіоналізація режиму атмосферних опадів західної частини України методом кластерного аналізу наявних даних спостережень метеостанцій.

Результати досліджень. Територія дослідження охоплювала західну частину України в межах Львівської, Закарпатської, Івано-Франківської, Волинської, Тернопільської та Чернівецької областей. Період дослідження охоплював два роки: 1961 та 1970. Обидва роки характеризувались суттєво відмінним режимом випадання опадів та їхнім просторовим розподілом. 1961 рік відмічений суттєво меншими від середніх річними кількостями опадів, з переважанням антициклонального режиму протягом більшості місяців. 1970 рік, навпаки, був більш вологим протягом усіх сезонів року, з частим проходженням циклонів і атмосферних фронтів, переважанням повітряних мас із заходу [1].

Зазначені відмінності між роками можна пояснити характером глобальних циркуляційних процесів. Ці процеси зазнають ряду періодичних та неперіодичних флуктуацій, викликаних складною динамікою глобальної кліматичної системи у її взаємодії зі Світовим океаном. Однією з найбільш значимих таких флуктуацій є Південна осциляція Ель-Ніньо (ENSO), яка найбільш чітко проявляється в південно-східній частині Тихого океану, проте має вплив на синоптичні процеси деінде. Добре описаним є вплив ENSO на синоптичні процеси в Європі. [6, 9]. Так, в періоди з феноменом Ель-Ніньо Ісландський мінімум над Північною Атлантикою послаблюється та зміщується на південь. Одночасно дещо послаблюється Азорський максимум. В цей період шляхи більшості циклонів пролягають над Середземномор'ям та Чорним морем. І навпаки, низькі (від'ємні) значення індексу ENSO свідчать про явище Ла-Нінья, яке супроводжується переміщенням Ісландського мінімуму на північ та послабленням циклональної циркуляції над Європою. За [5], у вологому 1970 р. індекс ENSO був вищим від 0 більшу частину року, тоді як у сухішому 1961 р., навпаки, більшу частину року він був від'ємним.

В нашому дослідженні використовувались дані про місячні суми опадів для метеостанції території дослідження (усього – 33 метеостанції для 1961 року та 35 метеостанцій для 1970 року), одержані з Метеорологічного щомісячника [1]. Для виділення кластерів використано метод к-середніх (з різною заданою кількістю кластерів) та метод ієрархічного кластерного аналізу, а для перевірки однорідності отриманих кластерів – метод дискримінантного аналізу. В цілому, результати застосування методів к-середніх та ієрархічного кластерного аналізу виявились подібними. На рис. 1 показано дендрограму, отриману для 1970 року.

Аналіз цієї дендрограми та результатів застосування методу к-середніх дає змогу виділити шість кластерів (рис. 2). Кластер 1 є найбільшим й включає метеостанції Волині, Поділля та Малого Полісся, а також рівнинного Закарпаття. Для нього був характерний досить рівномірний розподіл кількостей опадів за порами року, з найменшими, серед інших кластерів, кількостями опадів у зимовий час (табл. 1).

Кластер 2 включає метеостанції західної частини Українських Карпат, в основному на висоті менше 700 м (станція Плай, розташована вище, дещо вирізняється за режимом опадів серед інших членів кластеру);

Кластер 3 включає метеостанції північно-східної частини Українських Карпат і Передкарпаття. Він характеризувався нерівномірним розподілом опадів за сезонами, з великою кількістю опадів у травні та червні.



Рис. 1. Дендрограма ієрархічного кластерного аналізу розподілу атмосферних опадів на метеостанціях заходу України за 1970 рік.

Кластер 4 включає метеостанції смуги "зовнішнього" Передкарпаття, з найбільшою кількістю опадів у липні. Метеостанції Мостиська та Яворів за режимом опадів дещо наближені до кластеру 1.

Кластер 5 включає метеостанції смуги "внутрішнього" Передкарпаття. Режим опадів тут – проміжний між кластером 4 та "карпатськими кластерами";

Кластер 6 включає метеостанції південно-східної частини Українських Карпат. Цей кластер – найбільш внутрішньо неоднорідний поміж усіх кластерів. Він характеризується значно більшою кількістю зимових опадів, поміж інших кластерів.

Дендрограма для 1961 року (рис. 3) має дещо інший вигляд. Насамперед, більша довжина гілок (пор. з рис. 1) вказує на більш неоднорідні кластери та більші локальні відмінності режиму опадів між сусідніми метеостанціями. На

це вказує і результат дискримінантного аналізу: якщо для 1970 року значення критерію Фішера F (для виділених кластерів) становило 11,52, то для 1961 року – лише 6,68. Загальна кількість кластерів, що впевнено виділялись, у 1961 була більшою: 8 замість 6.

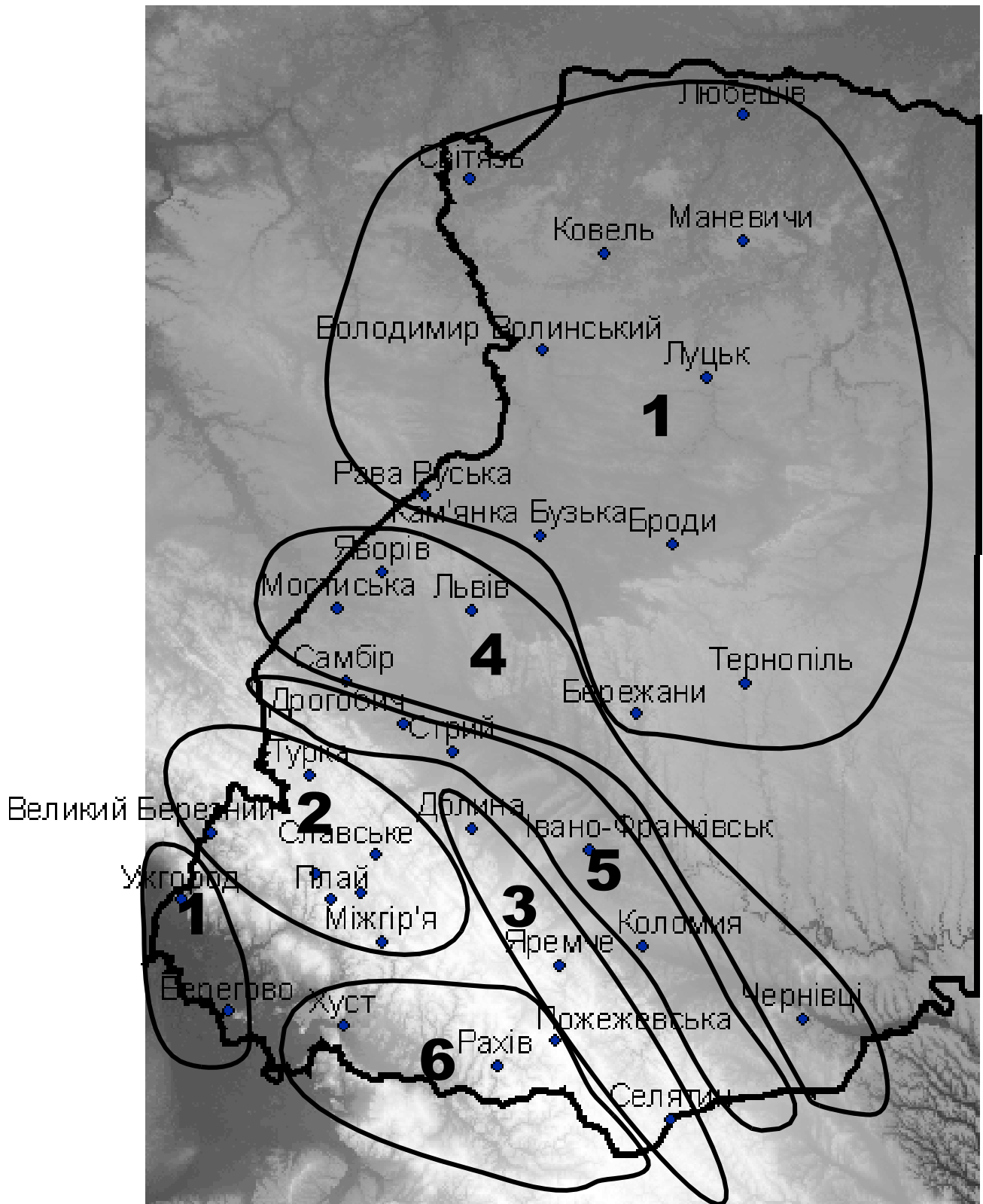


Рис. 2. Регіоналізація режиму опадів у 1970 р. Цифри – номери кластерів

Таблиця 1.

Усереднені за кластерами значення кількості опадів у 1970 р.

Кластери	Jan	Feb	March	Apr	May	June	July	Aug	Sept	Oct	Nov	Dec	Річна сума
1	57	46	41	45	100	68	87	51	69	41	64	66	736
2	74	90	109	107	139	143	187	82	70	64	75	103	1243
3	44	57	57	68	288	170	147	90	57	48	20	59	1106
4	58	44	42	54	117	87	156	55	58	41	43	57	811
5	46	43	50	50	195	104	124	65	54	38	36	63	869
6	103	181	156	111	275	171	125	82	125	106	99	139	1673
Усі	62	66	66	66	153	107	128	65	70	51	59	77	967

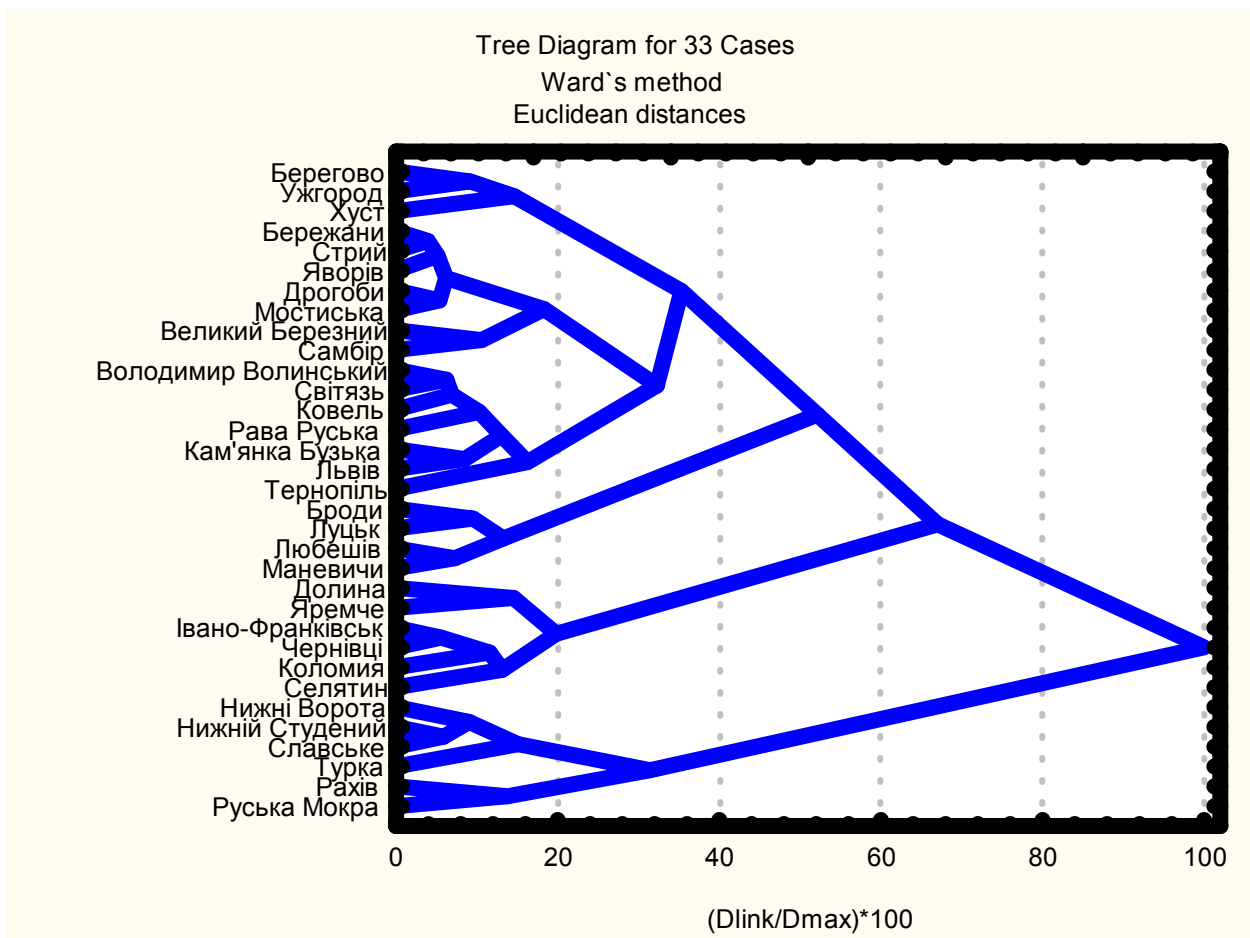


Рис. 3. Дендродіаграма ієрархічного кластерного аналізу розподілу атмосферних опадів на метеостанціях заходу України за 1961 рік.

Загальна картина просторового розподілу кластерів у 1961 році також суттєво відрізнялась від 1970 р. Насамперед, розміри кластерів були суттєво меншими (рис. 4). Якщо у 1970 р. кластери переважно розташовувались паралельно до простягання Карпат, то у 1961 р. більш помітним був меридіональний компонент диференціації режиму опадів.

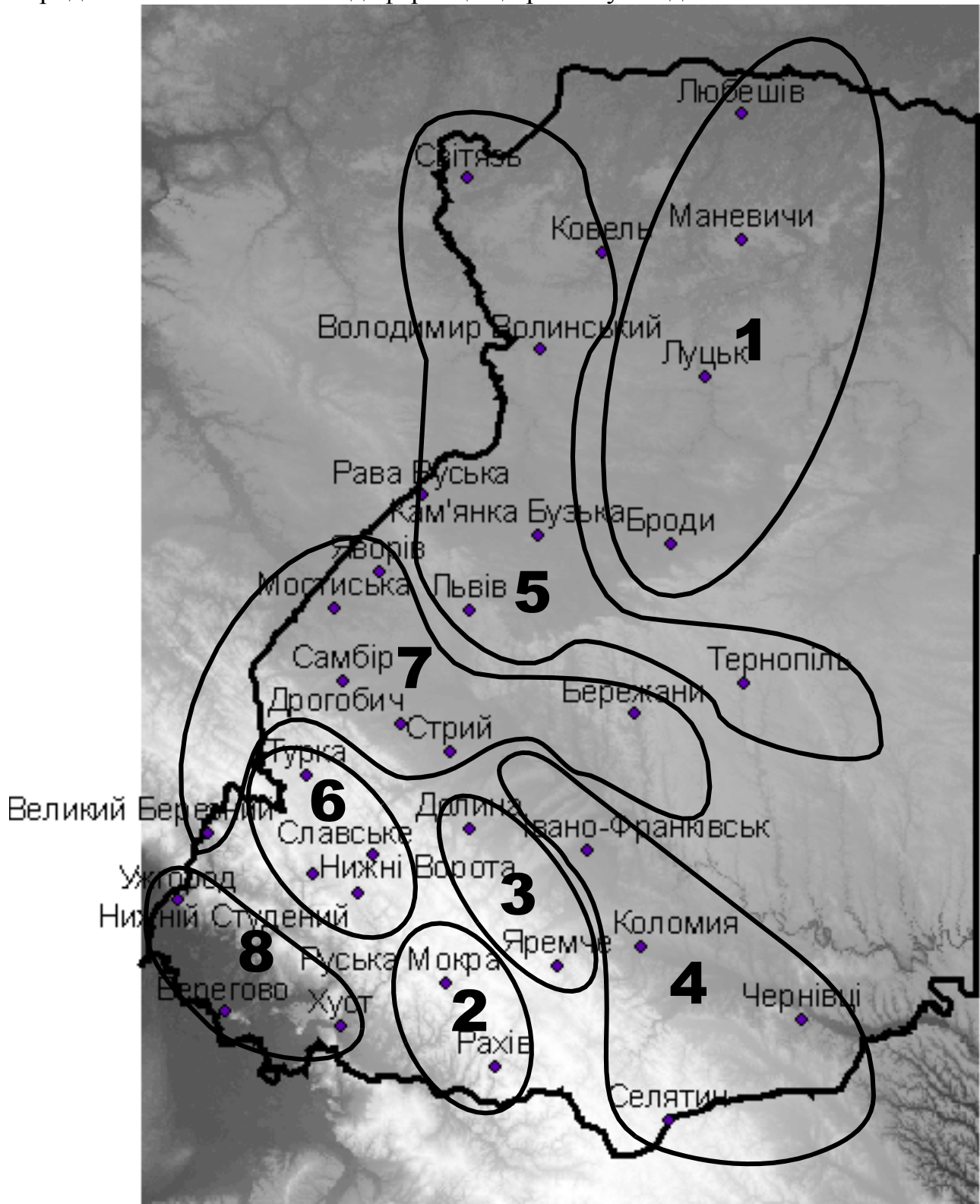


Рис. 4. Регіоналізація режиму опадів у 1961 р. Цифри – номери кластерів.

Таблиця 2.
Усереднені за кластерами значення кількості опадів у 1961 р.

Кластери	Jan	Feb	Mar	Apr	May	June	July	Aug	Sept	Oct	Nov	Dec	Річна сума
1	15	13	21	28	25	27	28	34	31	13	21	39	308
2	38	23	81	82	80	133	111	35	19	5	87	93	788
3	15	21	31	71	87	72	93	68	15	9	23	37	542
4	21	20	22	57	94	41	139	47	5	6	24	26	503
5	14	14	21	23	40	36	91	45	12	10	33	31	371
6	22	25	43	33	70	137	120	52	27	14	34	60	638
7	13	21	23	16	50	81	101	34	11	14	29	40	434
8	32	33	16	37	49	52	74	30	9	21	73	76	501
Усі	19	20	28	35	57	67	94	42	16	12	36	45	473

Висновки. Отже, кластерний аналіз метеорологічних даних метеостанцій певного регіону є ефективним засобом регіоналізації кліматичних полів. Аналіз даних метеостанцій заходу України для двох років, які суттєво відрізнялись за переважаючими синоптичними режимами, виявив, що в залежності від характеру глобальних циркуляційних процесів, які впливають на ці режими, результати регіоналізації можуть суттєво відрізнятись. Видається доцільним виконати аналогічний аналіз для ряду інших років, сформувавши репрезентативну вибірку та провівши детальний аналіз залежностей між розподілом кластерів та характером глобальних циркуляційних процесів в певні роки. Такий же аналіз можна провести й для усереднених даних, які характеризують кліматичну норму. Це дасть змогу отримати об'єктивну картину регіоналізації території за режимом опадів. Виконавши аналогічний аналіз щодо просторового розподілу інших метеорологічних елементів (температур повітря, сонячної радіації, режиму вітрів тощо), можна отримати об'єктивну, науково та статистично обґрунтовану картину кліматичного районування території.

1. *Метеорологический ежесемесечник.* –К., Выпуск 10, № 1-12, 1961, 1970.
2. *Baeriswyl P.A., Rebetez M.R.* Regionalization of precipitation in Switzerland by means of principal component analysis // *Theoretical and Applied Climatology.* – 1997. –Vol. 58, N. 1-2. –P. 31-41.
3. *Dinpajoooh Y., Fard A.F., Vahed M.M., Jahanbakhsh S., Mirnia M.K.* Selection of variables for the purpose of regionalization

of Iran's precipitation climate using multivariate methods // Iranian Journal of Agricultural Sciences. –2003. –Vol. 34, N. 4. –P.809-823. 4. Gong X., Richman M.B. On the application of cluster analysis to growing season precipitation data in North America East of the Rockies // J. Clim. –1995. –Vol. 8. –P. 897–931. 5. *International Comprehensive Ocean-Atmosphere DataSet (ICOADS)*, version 2.0, for 1800-2005 // Веб-ресурс <http://www.cdc.noaa.gov/coads/>. 6. Kiladis G.N., Diaz H.F. Global climatic anomalies associated with extremes in the Southern Oscillation // J.Clim. – 1989. –N. 2. –P. 1069-1090. 7. Munoz-Diaz D., Rodrigo F.S. Spatio-temporal patterns of seasonal rainfall in Spain (1912–2000) using cluster and principal component analysis: comparison // *Annales Geophysicae*. –2004. –Vol. 22. –P. 1435–1448. 8. Ward J.H. Hierarchical grouping to optimise an objective function // J. Amer. Stat. Assoc. –1963. –Vol. 58. –P. 236–244. 9. Wilby R. Evidence of ENSO in the synoptic climate of the British Isles since 1880 // *Weather*. –1993. –N. 48. –P. 234-239.

Мкртчян О.С., Шубер П.М. **Регіоналізація річного режиму атмосферних опадів західної частини України методом кластерного аналізу.**

В статті розглянуте застосування кластерного аналізу даних метеостанцій для регіоналізації режиму опадів – поділу території на регіони, однорідні за цим режимом. На основі аналізу місячних сум опадів метеостанцій заходу України за 1961 та 1970 роки виділено кластери метеостанцій, однорідних за режимом опадів. Виявлено, що кількість, розміри та межі таких кластерів відрізняються в залежності від переважаючих синоптичних режимів, зумовлених характером глобальних процесів атмосферної циркуляції.

Mkrtchian A., Shuber P. **Regionalization of annual precipitation regime in Western Ukraine using cluster analysis.**

The paper deals with the application of cluster analysis technique for the regionalization of the precipitation regime based on available meteorological data. The analysis of monthly precipitation data for Western Ukraine for years 1961 and 1970 allowed to distinguish clusters of stations with homogenous annual precipitation regime. It was found that the number of clusters, their sizes and boundaries vary greatly depending on predominant synoptic regimes, in turn dependent upon global atmospheric circulation patterns.