

СУЧАСНІ ГЕОГРАФІЧНІ ТА ЕКОЛОГІЧНІ ДОСЛІДЖЕННЯ ДОВКІЛЛЯ

УДК 912.681

С. В. КОСТРИКОВ, д-р геогр. наук, проф.
Харківський національний університет імені В. Н. Каразіна
майдан Свободи, 4, м. Харків, 61022
sergiy.kostrikov@geocloud.com.ua

ВИБІРКИ ТА ЗАПИТИ ЯК БАЗОВІ ГІС-ОПЕРАЦІЇ ПРИ ВИРІШЕННІ ГЕОЕКОЛОГІЧНИХ ЗАДАЧ

Подается загальна концепція базових функціонально-аналітичних операцій в середовищі геоінформаційних систем. Обговорюються деякі особливості окремих ГІС-операцій, які доцільно використовувати в предметній галузі геоєкології. На прикладах вирішення регіональних геоєкологічних задач доводиться, що вибірки та запити по інформації, що знаходиться в БД ГІС, є ефективним засобом, який може бути успішно застосований як до просторових, так і до атрибутивних даних, щоб, зокрема, сприяти досягненню цілей територіального природоохоронного менеджменту.

Наводяться релевантні приклади ГІС-операцій з посиланням на графічний інтерфейс користувача програмного забезпечення відомої ГІС-платформи.

Ключові слова: функціонально-аналітична операція ГІС, операції над просторовими та непросторовими даними, вибірки та запити, інтерактивна вибірка, Булева алгебра логіки

Kostrikov S. V. SAMPLING AND REQUESTS AS BASIC GIS OPERATION OF DECISION GEO-ECOLOGICAL PROBLEMS

The paper represents the general concept of the GIS analytical capabilities provided through basic GIS-operations. Some particular procedures of these operations are described and discussed due to the geoeological subject area. Few examples of the regional environmental tasks are provided with their solutions through GIS-operations. It is proved that data selections and queries to both spatial and non-spatial information are the effective method, which can be performed for the areal environmental management.

The paper provides some relevant interface examples from a well-known GIS-platform.

Key words: a functional-analytical GIS-operation, GIS-operations on spatial and non-spatial data, data selections and queries, interactive data selection, Boolean logic

Костриков С. В. ВЫБОРКИ И ЗАПРОСЫ КАК БАЗОВЫЕ ГИС-ОПЕРАЦИИ ПРИ РЕШЕНИИ ГЕОЭКОЛОГИЧЕСКИХ ЗАДАЧ

Излагается общая концепция базовых функционально-аналитических операций в среде геоинформационных систем. Обсуждаются некоторые особенности отдельных ГИС-операций, которые целесообразно использовать в предметной отрасли геоэкологии. На примерах решения региональных геоэкологических задач обосновывается, что выборки и запросы по информации, которая находится в БД ГИС, являются эффективным инструментом, который может быть успешно применен как к пространственным, так и к атрибутивным данным. В частности, для того, чтобы способствовать достижению целей территориального природоохранного менеджмента.

Приводятся релевантные примеры ГИС-операций с ссылкой на графический интерфейс пользователя программного обеспечения известной ГИС-платформы.

Ключевые слова: функционально-аналитическая операция ГИС, операции над пространственными и непространственными данными, выборки и запросы, интерактивная выборка, Булева алгебра логики

Вступ

В нещодавно опублікованому авторському навчальному посібнику з геоінформаційних систем (ГІС) окремо підкреслюється достатньо відома істина – найбільш характерною рисою сучасного суспільства є процеси -

його *інформатизації* – тотального проникнення програмних і апаратних засобів у всі сфери суспільно-історичної практики людства [1]. Зараз є загальнозрозумілим, що на сучасному етапі розвитку методів збору, обробки та аналізу *просторово-координованої інформації (ПКІ)* зростає роль ГІС, як універса-

льного інструменту географічного та екологічного дослідження, причетного до цих двох предметних галузей. Оскільки, за своєю суттю *географічна модель ПКІ* визначає тезаурус опису і пояснення об'єктів і процесів, які знаходяться в межах географічної оболонки, а *екологічна модель ПКІ* – відповідні дефініції щодо глобальної соціоекосистеми, то важко переоцінити значення подібних моделей для ключового змісту університетської освіти щодо географічних і екологічних дисциплін. Це, зрозуміло, має відноситися до теоретичного аспекту навчального процесу підготовки фахівців із географії та екології. В цей саме час, в прикладному аспекті треба мати на увазі, що географічні та екологічні моделі ПКІ є тими засадами, на яких будуються буквально всі сучасні геоінформаційні платформи.

Через поглиблений огляд фундаментальних публікацій щодо функціональності ГІС-платформ автором статті свого часу був зроблений висновок щодо неповноти класифікації *базових операцій* в ГІС, які стосуються більшості дослідницьких задач в предметних галузях геоєкології, геоморфології та гідрології. На підставі чого була запропонована наступна вдосконалена класифікація [2] (рис. 1):

Результати дослідження

Відповідно до вдосконаленої класифікації, що створена на підставі її першої версії (рис. 1), базові предметні ГІС-операції, які стисло характеризуються в цьому розділі, доцільно поділити на наступні класи:

- Операції, які впроваджуються в базах даних ГІС (в базах геоданих) над просторовими та непросторовими даними – *вибірки та запити до даних* (інтерактивна вибірка, побудова довідкової таблиці, створення числових порогів для вибірки по кількісним даним; застосування Булевої алгебри логіки); *перерахунок*;

- Операції, які виконуються на окремому шарі просторових даних – *конвертація одних структур і форматів даних в інші; вимірювання* (атрибутивні вимірювання, вимірювання розміру, фрактальна розмірність, індексування зразків та шаблонів географічних сутностей); *просторові агрегація та генералізація; операції «ковзного вікна»; аналіз безпосереднього сусідства й накладання; створення буферних зон; просторова фрагментація*;

- 1) операції із базами геоданих – базові функціонально-аналітичні операції ГІС, які передбачають процедури збереження, обробки й аналізу як просторової, так і непросторової інформації; 2) операції, які забезпечують процедури обробки та аналізу окремих шарів просторової інформації; 3) операції, які передбачають геообробку та аналіз композитних (складних) шарів просторової інформації.

Враховуючи різницю між растровими та векторними моделями, треба мати на увазі, що існують певні методичні відмінності при впровадженні ГІС-операцій під час роботи з даними одного чи іншого типу, хоча переважна більшість цих операцій схожим чином застосовується для даних обох моделей [5]. У випадку окремого розгляду растрових та векторних структур даних слід мати на увазі, що растрові структури передбачають кількісні та якісні *процедури геообробки* у той час, коли векторні структури – *процедури просторового та атрибутивного аналізу*.

Метою статті є узагальнене подання відповідно уточненої авторської класифікації основних класів тих базових операцій в середовищі ГІС, які переважно використовуються при вирішенні геоєкологічних задач.

- Операції, які виконуються на декількох шарах просторових даних – *геометрична трансформація* (реєстрація через «абсолютну позицію», реєстрація через «відносну позицію», об'єднання та узгодження країв); *вирізки; операції оверлею* (графічний оверлей, логічний оверлей, арифметичний оверлей);

- *Функціональні операції геообробки та геовізуалізації* – на відміну від усіх попередніх базових ГІС-операцій є комплексними сукупностями процедур і операцій в середовищі геоінформаційної платформи, двома з трьох головних складових її функціональності – база геоданих, геообробка та геовізуалізація [2].

Детальніше зупинимося лише на першій групі вказаних ГІС-операцій, яка, на нашу думку, є найбільш функціональною при геоєкологічних дослідженнях.

Операції в БД ГІС над просторовими та непросторовими даними – найбільш широкий клас ГІС-операцій, на який існує попит в предметних галузях географії та екології.

Програмне забезпечення ГІС, подібно іншим СУБД, може бути застосоване для

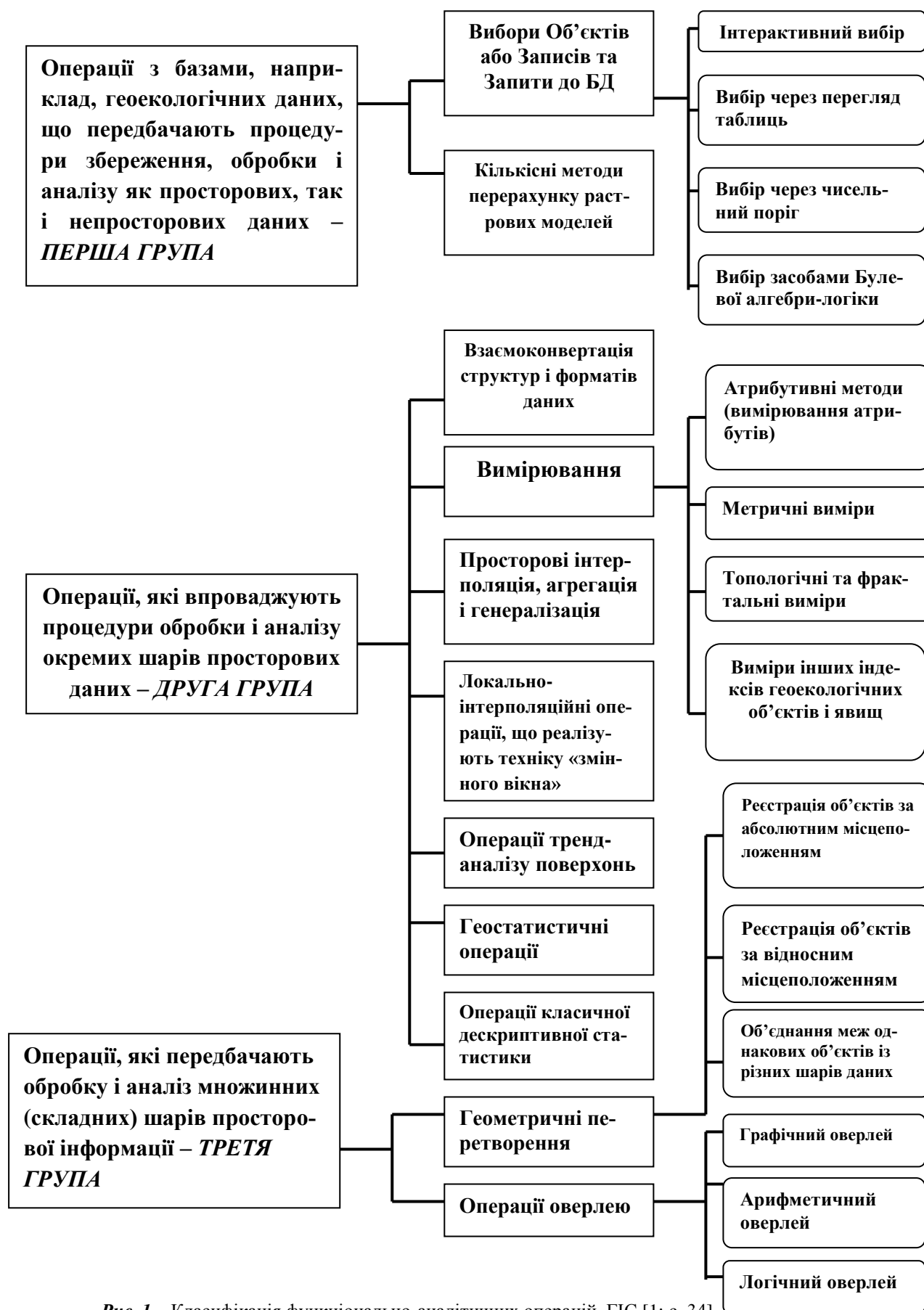


Рис. 1 – Класифікація функціонально-аналітичних операцій ГІС [1; с. 34]

вибірки по масиву даних і для побудови запитів до даних, які задовольняють певним критеріям. Незважаючи на схожість концептуального підходу щодо вибірок та запитів, який застосовується в «звичайних СУБД» і в ГІС, в рамках подібних операцій непотрібні масиви «зайвих даних» відфільтровуються саме тими ГІС-засобами, через які виконуються ключові рішення в рамках другої та третьої груп операцій, зазначених на ілюстрації вище (див. рис. 1). Тобто йдеться про «операції на окремому шарі просторових даних» та про «операції, які виконуються на декількох шарах просторових даних».

Вибірки та запити: підсумовування таблиці та інтерактивна вибірка. Наступний приклад складений на підставі учбових даних для ГІС-платформи ArcView 3.X [6].

Припустимо, що фахівець із стану природного довкілля працює в Африці в рамках програми допомоги ООН для країн «третього світу». За допомогою ГІС він має спрогнозувати для регіону Марсабіт (Кенія), які території в найбільшій мірі є вразливими щодо можливих засух. Це має запобігти загибелі худоби і, відповідно – розоренню місцевих фермерів. В якості первинних даних фахівець має *ізолійну карту*, яка відбиває розподіл річної суми опадів у цьому регіоні. Території з річною сумою опадів, меншою ніж 300 мм, треба вважати такими, які є під загрозою засухи. Території, які відповідають цьому критерію, вибираються з *атрибутивної таблиці* шару *Rainfall*, який є сукупністю полігонів (регіонів), границя яких проходить по ізолінії певної річної кількості опадів (рис. 2).



Рис. 2 – Просторова і атрибутивна інформація в БД ГІС ArcView 3.x

Місцеположення певного полігону (екорегіону) є *просторовою інформацією*, а сума опадів, що асоціюється з кожним полігоном – *непросторовою*. Всі дії фахівця з цими даними при наповненні БД ГІС підпадають під широку категорію операцій і процедур з просторовими та непросторовими даними, ключовими з яких є, що вже підкреслювалося вище – *вибірки та запити*.

Всі екорегіони, до яких прив'язані значення опадів, мають бути поділені на дві категорії: 1) *засушливі* (*dry* – англ.), для яких ці значення менше 300 мм та 2) *помірні* (*medium* – англ.) із значеннями більше 300 мм (рис. 2). Фахівець обирає необхідні (*dry*) екорегіони – просторові об'єкти – че-

рез вибірку записів – атрибутів відповідно до вказаного кількісного критерію. Можна буде визначити загальну площу, яка може бути вражена засухою, а потім поєднати всі потенційно засушливі регіони в один об'єкт.

Стандартним засобом ГІС-платформи для базової ГІС-операції запитів є діалог *Виробника Запитів*, через який обираються всі записи в таблиці, що задовольняють обраному критерію суми опадів (рис. 3).

За результатами запиту буде отримано наступне статистичне повідомлення:

Бачимо, що засушлива площа складає майже 42 тисячі кв. км, тобто її складають регіони, для яких сума опадів буде меншою

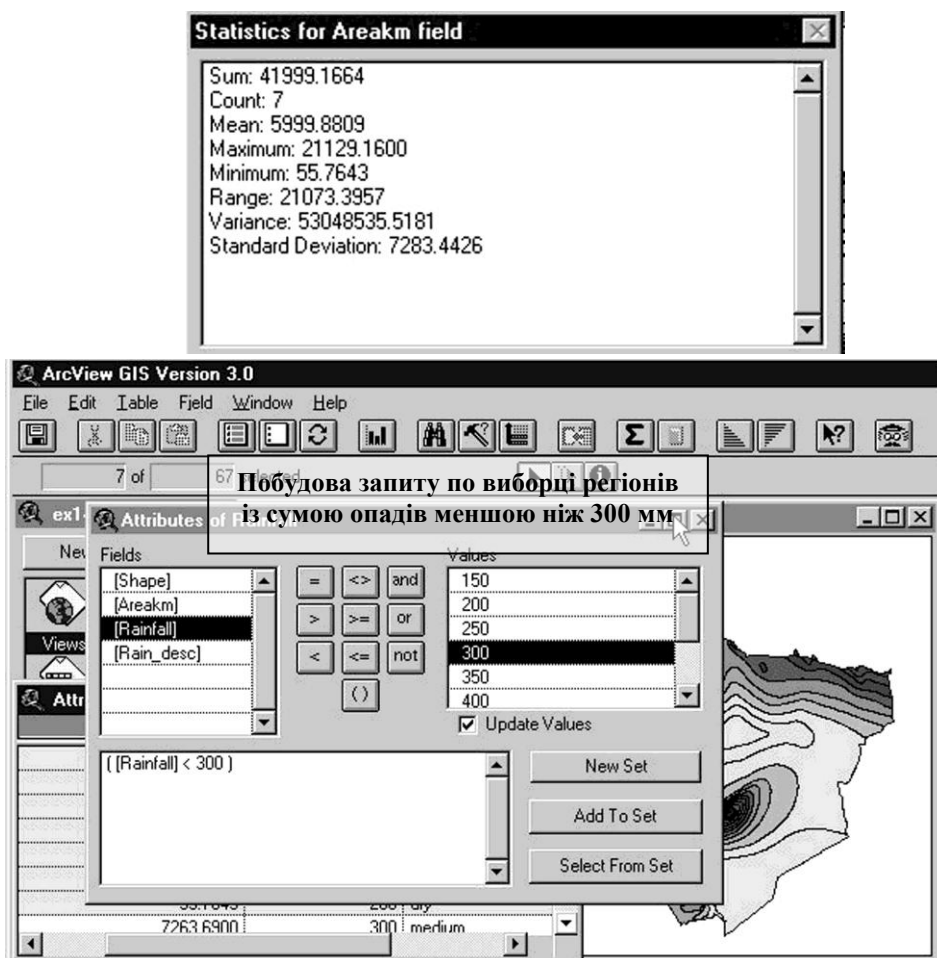


Рис. 3 – Діалог Виробник Запитів у середовищі ГІС ArcView 3.x

за 300 мм. Крім сумарної площі наводяться основні статистичні показники – характеристики центру групування і розсіювання.

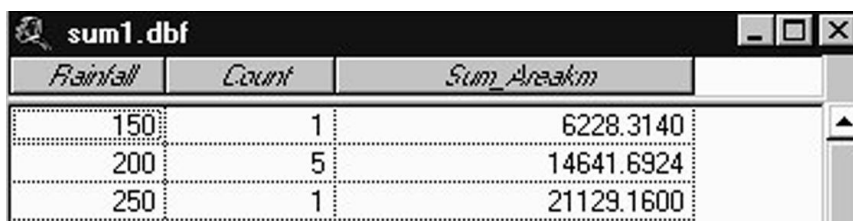
Після вибірки через побудований запит виконується така процедура як Підсумовування Таблиці. Коли ГІС впроваджує цю процедуру, створюється нова таблиця, яка містить щонайменш два поля (стовпчик): одне, для унікальних значень, що, власне, і підсумовується; інше поле вказує скільки разів в таблиці було знайдено кожне унікальне значення. Припустимо, ГІС-користувач відкрив атрибутивну таблицю, яка складається з 10 записів (стрічок). Поле (стовпчик) таблиці, яке називається *Тип* вказує на те, що подає кожен запис – наприклад, фабрику, або ферму. Коли ГІС-користувач впроваджує процедуру Підсумовування Таблиці по цьому полю, він в результаті отримає похідну таблицю з двома записами (стрічками) тільки в окремих стовпчиках: значення «фабрика» та «3»; зна-

чення «ферма» та «7». ГІС-карта території, де розташовані ці об'єкти, візуалізувала би 10 дискретних сутностей двох типів із таким саме розподілом – 3 проти 7.

Здійснення процедури Підсумовування Таблиці в регіональному прикладі, що розглядається, має, як результат рис.4.

Таким чином, через операції вибірки і запиту до атрибутивної таблиці встановлено, що серед семи полігонів, які задовольняють значенню <300 мм річної суми опадів, два опиняються в класах крайніх значень (150 і 250, відповідно) і п'ять – в середньому класі (200). Поле *Sum_Areakm* демонструє фахівцю, котрий досліджує район Марсабіт, яка за розміром площа із тих, що є потенційно засушливими, отримає лише 150 мм опадів на рік, яка – 200 мм, а яка – 250 мм.

Побудова довідкової таблиці та вибірка через числові порогови. Підтип базових ГІС-операцій вибірок та запитів, через



Rainfall	Count	Sum_Areakm
150	1	6228.3140
200	5	14641.6924
250	1	21129.1600

Рис. 4 – Вигляд підсумкової таблиці

який зроблена просторова класифікація унікальних значень по полю *Rainfall* (дощ) атрибутивної таблиці, є так званою *інтерактивною вибіркою*. Такий прийом застосовується для простих вибірок і, відповідно, для просторової класифікації по відносно невеликим таблицям атрибутів, коли число записів (стрічок) таке, що дозволяє інтерактивно створювати просторові класи об'єктів.

При складних вибірках або при простих, але таких що регулярно повторюються, застосовується інший підтип ГІС-операцій із типу операцій вибірок та запитів, який визначається як *побудова довідкової таблиці*. Такий прийом не розділяє первинні та похідні дані в атрибутивних таблицях і саме тому дозволяє в деяких випадках оперативно впроваджувати більш узагальнені схеми просторової класифікації. Наприклад, через побудову довідкової таблиці щодо району Марсабіт (щодо прикладу, наведеному вище, і відповідній класифікації) можна всі полігони (екорегіони) поєднати в три загальні класи: *засушливі*, *помірні* та *вологі*. Таким чином, буде виконана процедура *злиття ГІС-об'єктів*, яку ми дещо детальніше розглянемо далі в тексті посібника при розгляді відповідного типу ГІС-операцій.

Щодо впровадження операції «побудова довідкової таблиці» треба зробити важливе зауваження. Оскільки ця операція часто виконується для *просторової перекласифікації*, треба пам'ятати, що хоча перекласифікація і може генерувати нові дані, ці дані ні в якому разі не можуть бути більш детальними ніж інформація, отримана на кроці первинної просторової класифікації. Вторинна БД ГІС, що була побудована на підставі перекласифікації просторових даних із первинної БД ГІС, яка мала, припустимо, вісім класів, повинна мати також вісім класів даних *або менше*. Дев'ятий клас у цьому випадку (крім, зрозуміло, «пустих» класів як-то *Невідомі* або *Інші*) не може бути доданий без отримання додаткових емпіричних даних.

Таким чином, дуже важливо при роботі з БД ГІС починати з детальної класифікації через операцію «*інтерактивна вибірка*» та застосовувати прийом «побудова довідкової таблиці» для подальшої генералізації через перекласифікацію, якщо такі дії є необхідними.

Створення *числових порогів* для вибірки по кількісним даним переважно застосовується в геоекологічних ГІС-застосуваннях. Найпоширеніший приклад – просторова класифікація будь-якої, природно виокремленої території, наприклад, водозбірного басейну по таксонам розповсюдження вздовж річкового русла безхребетних тварин класу комах – *весьнянок* (*Plecoptera*), *поденок* (*Ephemeroptera*) та «струмочників» (*Trichoptera*). На підставі зібраних польових спостережень складалася, в даному випадку, вузькопредметна «екологічна» БД ГІС, яка містила *ЕРТ-таксони* (*таксономічні групи* по перших літерах латинських назв цих безхребетних) на двох водозборах річок штату Мічиган (США) – ділянки переважного мешкання тих або інших видів безхребетних [4] (рис. 5).

Висока якість природного довкілля в межах водозбірного басейну визначалася *індикативним методом* на підставі показника *багатства безхребетних* – *R* (*Richness* – англ.). Цей показник є похідною величиною від кількості *агрегованих ЕРТ-таксономічних груп* на ділянку площі вздовж русла річки і взагалі по водозбору. Причому, ця величина залежить від *співвідношення* між собою цих різних видів комах, що визначається за наступною простою формулою [4]:

$$\text{Агреговані ЕРТ-таксони} = \text{Таксони Ephemeroptera} + \text{таксони Plecoptera} + \text{таксони Trichoptera}$$

Показник *R* непрямо залежить від значення агрегованих таксонів, і останнє (кількість ЕРТ-таксонів) навіть скоріше за

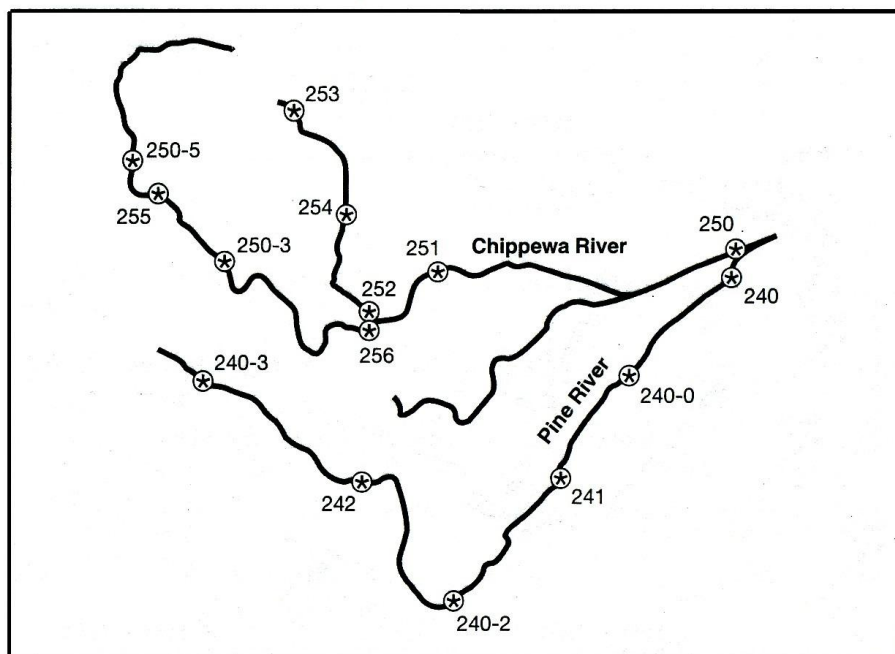


Рис. 5 – Точки спостереження по водозборам двох річок (штат Мічиган, США) щодо наповнення «екологічної» БД ГІС – дослідження розповсюдження безхребетних тварин з класу комах; по кожній точці прописані ID ділянки вздовж русла, наприклад 251 [4, с. 51]

показник R використовується для оцінки якості довкілля водозбору. Взагалі: велика кількість ЕРТ-таксонів – висока якість довкілля; низька кількість – низька якість. Таким чином і встановлюється прос-

торова класифікація через вибірку по числовому порозу, однак, частіше за все, встановлюється багаторазовий числовий поріг вибірки, як на наступній ілюстрації (рис. 6):

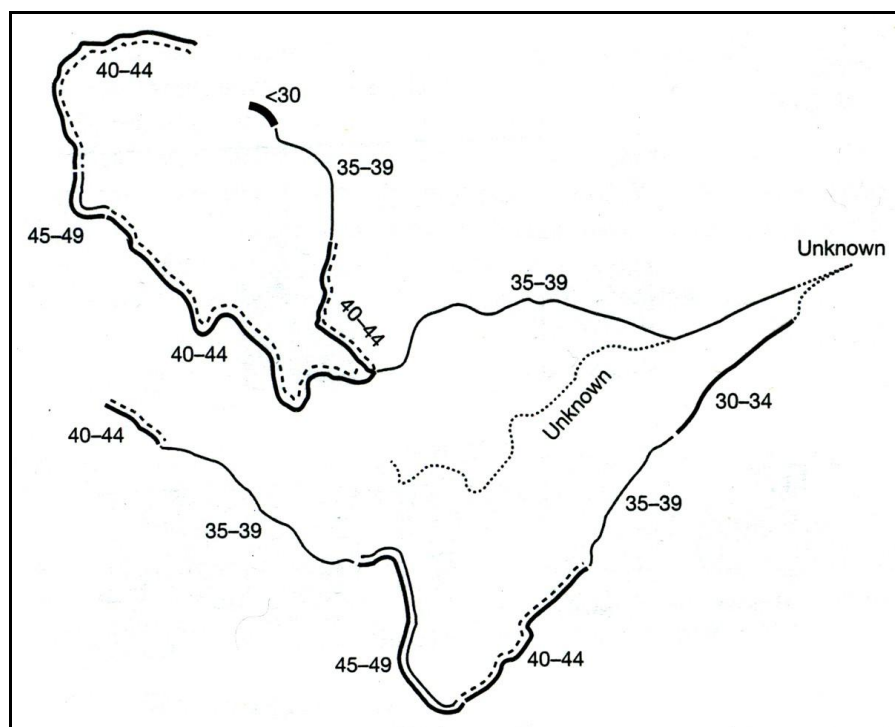


Рис. 6 – Ділянки русел по водозборах двох річок (див. рис. 4), класифіковані за показником R через вибірку по числових порогах [4, с. 75]

Показник R по басейнах цих двох річок коливається в межах 31-49, і вибірка через числові пороги на підставі агрегованих *EPT*-таксонів фактично впроваджується, щоб районувати ділянки русел по *характеристиці біологічного різноманіття*. Для цього і встановлюється багаторазовий числовий поріг вибірки що, в свою чергу, дозволяє зробити просторову класифікацію окремих ділянок русла по обраному екологічному показнику (рис. 6). Окремі ділянки русел просторово класифікуються по числових порогах між такими значеннями R як: < 30; 30-34; 35-39; 40-44; 45-49. В даному

випадку нижні границі просторових класів (<30, 30, 35, 40, 45) є критеріями вибірки по числовому порогу.

Вибірki через Булеву алгебру логіки. Такий прийом часто використовується при формуванні вибірок із сукупностей географічних даних та запитів до них. Його відносять до *локальних операцій картографічної алгебри* [3]. В предметній галузі ГІС цей розділ математики використовує наступні оператори для впровадження операцій вибірки по двох і більше сукупностях даних, відповідно так званій «діаграмі Венна» (рис. 7):

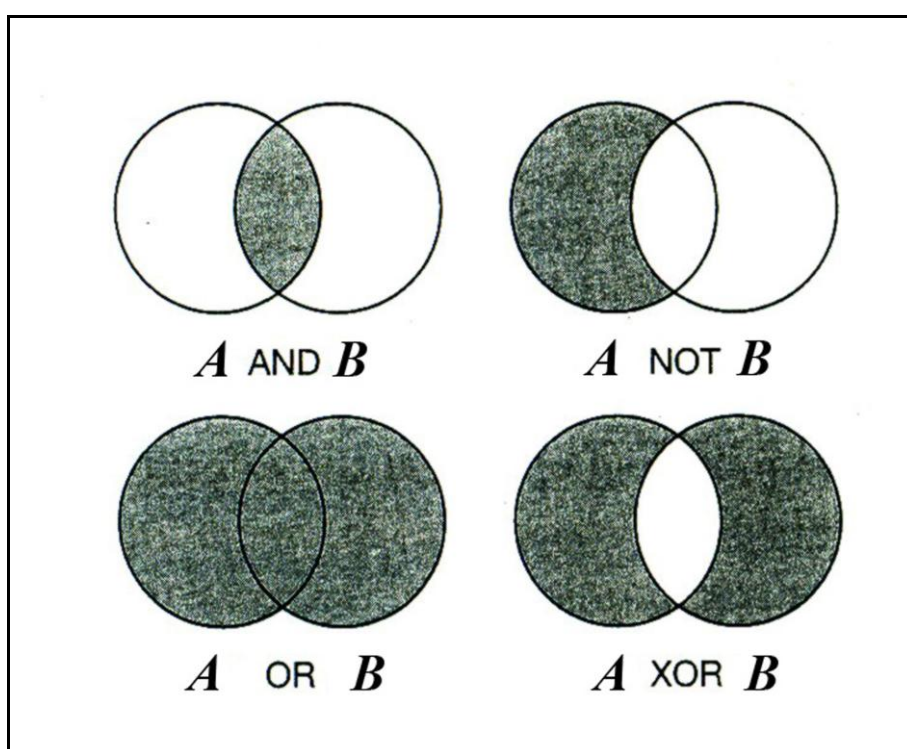


Рис. 7 – Діаграма Венна, що подає операції Булевої алгебри логіки, які можуть застосовуватися для просторової вибірки

Базовими Булевими операторами є наступні:

AND – *логічне множення* (логічне *і*), яке у випадку просторової вибірки через Булеву алгебру, відбиває *перетинання* двох територій (рис. 7);

OR – *логічне складання* (логічне *або*), яке у випадку просторової вибірки через Булеву алгебру, відбиває *поєднання* двох територій;

NOT – *логічне заперечення* (логічне *ні*), яке у випадку просторової вибірки че-

рез Булеву алгебру, відбиває *поєднання* двох територій;

OR – *логічне виключення* (логічне *або/або*), яке у випадку просторової вибірки через Булеву алгебру, відбиває *незбіг* двох територій.

При вибірках через логічні оператори із атрибутивних таблиць ГІС завжди є задіяні *два* (або більше) *атрибути*. Якщо в якості прикладу, звернемося до табличних даних, що наводяться в літературному джерелі, на яке ми посилалися вище [4], то можна обрати дві наступні атрибутивні характери-

стики. Останні відносяться до досліджуваної території, яка наводиться на двох ілюстраціях вище (рис. 5, 6).

Таким чином, наприклад, якщо атрибутивною характеристикою **A** є «біологічний рід = *Cladopelma* (лат.)», а атрибутивною характеристикою **B** є «число агрегованих *EPT*-таксонів ≥ 12 », тоді ми отримаємо наступні результати впровадження Булевих операцій:

A AND B – ділянки русел, де домінує біологічний рід *Cladopelma*, і для яких виконується умова «*EPT*-таксонів ≥ 12 »;

A NOT B – ділянки русел, де домінує біологічний рід *Cladopelma*, і для яких виконується умова «*EPT*-таксонів < 12 »;

A OR B – ділянки русел, де домінує біологічний рід *Cladopelma*, або для яких виконується умова «*EPT*-таксонів ≥ 12 »; обидві умови можуть виконуватися разом;

A NOT B – ділянки русел, де або домінує біологічний рід *Cladopelma*, або для яких виконується умова «*EPT*-таксонів ≥ 12 »; має виконуватися лише одна із цих двох умов.

Наступна таблиця подає результати вибірки із застосуванням Булевої алгебри логіки через вказані вище умови по двох атрибутивних характеристиках на території двох водозборів (рис. 5, 6).

Порівняння результатів просторової класифікації по вибірках через застосування числових порогів (рис. 6), з одного боку, і через застосування операторів Булевої алгебри логіки (табл.), з іншого, надає фахівцю-екологу додаткові важелі для прийняття зважених рішень у природоохоронному менеджменті.

Таблиця

Вибірка через оператори Булевої алгебри логіки на території двох водозборів (див. рис. 5)

Булеві операції	ІД точки спостереження-ділянки вздовж русла, на якій виконується умова даної Булевої операції
A AND B	250-5, 255
A NOT B	240, 240-0, 241
A OR B	240-2, 242, 250, 255
A XOR B	240, 240-2, 250-3, 254, 256

Висновок

На прикладах вирішення регіональних геоecологічних задач доводиться, що вибірки та запити по даних, які знаходяться в БД ГІС, є ефективною функціонально-аналітичною операцією ГІС. Дана базова ГІС-

операція може бути успішно застосована як до просторової, так і до атрибутивної інформації та, наприклад, сприяти досягненню цілей територіального природоохоронного менеджменту.

Література

1. Бережной В. А. Работа в среде ГИС-платформы *MAPINFO* - компьютерный практикум / В. А. Бережной, С. В. Костриков. – Х.: ХНУ ім. В.Н. Каразіна, 2015. – 108 с.
2. Костриков С. В. Геоінформаційне моделювання природно-антропогенного довкілля: наукова монографія / С.В. Костриков. – Х.: Вид-во ХНУ ім. В.Н. Каразіна, 2014. – 484 с.
3. Світличний О. О. Основи геоінформатики: навчальний посібник / О. О. Світличний, С. В. Плотницький. – Суми : Університетська книга, 2006. – 295 с.

4. Arthur J. Evaluation of watershed quality in the Saginaw River Basin / J. Arthur, T. Roush, J. Thompson // US Environmental Protection Agency Publication EPA/600/R-95/153. – Washington, DC: EPA Office, 1996. – 183 p.
5. Bailey J. ArcCatalog: Руководство пользователя / J. Bailey, A.Vienneau. - - М.: Дата+, 2006. - 265 с.
6. *Getting to Know ArcView GIS*. 3rd Edition – Redlands, CA: ESRI Press, 1999. – 411 p.

Надійшла до редколегії 29.04.2015