

МІНІСТЕРСТВО ОСВІТИ І НАУКИ УКРАЇНИ
ОДЕСЬКИЙ ДЕРЖАВНИЙ ЕКОЛОГІЧНИЙ УНІВЕРСИТЕТ

Факультет комп'ютерних наук,
управління та адміністрування
Кафедра інформаційних техно-
логій

МАГІСТЕРСЬКА КВАЛІФІКАЦІЙНА РОБОТА

на тему: Алгоритми просторового аналізу для вирішення
завдань геомаркетингу

Виконав студент 2 курсу групи МІС-21
спеціальності 122 Комп'ютерні науки
Кузніченко Андрій Дмитрович

Керівник д.т.н., професор
Казакова Надія Феліксівна

Рецензент к.т.н., доцент
Гнатовська Ганна Арнольдівна

Одеса 2022

АНОТАЦІЯ

на магістерську роботу

«Алгоритми просторового аналізу для вирішення завдань геомаркетингу»

Магістерська робота присвячена актуальній темі створення просторової систем підтримки прийняття рішень у сфері геомаркетингу, шляхом інтеграції методів мультикритеріального аналізу рішень в геоінформаційні системи.

Метою магістерської роботи є розроблення моделі для вирішення проблеми пошуку придатного місця розташування об'єкту громадського харчування (кав'ярні) на території міста Одеса, а саме дослідження методів отримання і обробки просторової інформації для застосування їх в моделі, що розробляється, а також підготовка вихідних просторових даних для моделювання.

Об'єкт дослідження – просторова система підтримки прийняття рішень по розташуванню об'єкту громадського харчування.

Предмет дослідження – методологія отримання і обробки просторової інформації для застосування її в просторовій моделі прийняття рішень.

Методи дослідження: аналіз придатності, методи просторового моделювання, аналізу ієрархій, нечіткої логіки, зваженої лінійної комбінації, а також методи зваженого накладання растрових даних і алгебра карт.

Результати дослідження: описані методологічні основи мультикритеріальної моделі прийняття рішень щодо визначення оптимального місця розташування об'єкту громадського харчування (кав'ярні) на території міста Одеса. Надані рекомендації щодо підготовки вихідних просторових даних для реалізації етапів моделі: обґрунтування вибору критеріїв та альтернатив, розроблення ієрархічної структури критеріїв, нормалізація та стандартизація критеріїв. Створено карти критеріїв та шари обмежувальних факторів. Побудована програмна реалізація запропонованих етапів в пакеті ArcGIS 10.2.

Ключові слова: ГІС, система підтримки рішень, мультикритеріальний аналіз рішень, метод аналізу ієрархій, нечітка логіка.

SUMMARY

Master's work on the topic "Algorithms of Spatial Analysis for Solving Geomarketing Tasks".

The master's work is devoted to the actual topic of creating spatial decision support systems, by integrating methods of multi-criteria analysis of solutions in geographic information systems.

The purpose of the master's work is the development of a model for solving the problem of finding a suitable location for a public catering facility (cafe) in the city of Odessa. The study of methods for obtaining and processing spatial information for their application in the developed model, as well as the preparation of output spatial data for modeling.

Object of study is a spatial decision support system based on the location of a public catering facility.

Subject of study is methodology for obtaining and processing spatial information for its application in the spatial decision-making model.

Research methods are suitability analysis, methods of spatial modeling, analysis of hierarchies, fuzzy logic, weighted linear combinations, as well as methods of weighted overlay of raster data and algebra maps.

Research results have described the methodological foundations of the multicriteria model of decision making on determining the optimal location of a public catering object (cafe) in the city of Odessa. The given recommendations for preparation of outgoing spatial data for realization of stages of the model: the substantiation of choice of criteria and alternatives, development of hierarchical structure of criteria, normalization and standardization of criteria. Criteria and layers of bounding factors have been created. The program implementation of the proposed stages in the ArcGIS 10.2 package was built.

Keywords: GIS, decision support system, multicriterial decision analysis, hierarchy analysis method, fuzzy logic.

ЗМІСТ

ПЕРЕЛІК УМОВНИХ ПОЗНАЧЕНЬ, СКОРОЧЕНЬ І ТЕРМІНІВ	8
ВСТУП	9
1 ДОСЛІДЖЕННЯ ЗАСОБІВ ГІС ДЛЯ ПІДТРИМКИ ПРИЙНЯТТЯ ПРОСТОРОВИХ РІШЕНЬ	11
1.1 Географічні інформаційні системи	11
1.1.1 Моделі даних	12
1.1.2 Приклади просторового аналізу	14
1.2 Порівняльний аналіз сучасних ГІС пакетів.....	18
1.2.1 Програмний комплекс ArcGIS Desktop	18
1.2.2 Пакет ГІС MapInfo Professional	21
1.2.3 Пакет QGIS і інші відкриті настільні ГІС	23
2 АНАЛІЗ ПІДХОДІВ І МЕТОДІВ ВИРІШЕННЯ ЗАДАЧІ	26
2.1 Огляд методів Data Mining.....	26
2.2.1 Зважена сума (Weighted Sum).....	29
2.2.2 Метод аналізу ієрархій (АНР)	30
2.2.3 Нечітка логіка.....	30
2.2.4 Метод зваженого лінійного комбінування.....	35
3 ПОБУДОВА МОДЕЛІ.....	39
3.1 Постановка завдання.....	39
3.2 Вихідні дані моделювання	40
3.3 Визначення критеріїв.....	42
3.4 Побудова обмежуючого шару	43
3.5 Побудова шарів критеріїв та їх стандартизація	44
3.6 Визначення ваги критеріїв методом аналізу ієрархій	49
3.6.1 Побудова ієрархії критеріїв	51
3.6.2 Побудова матриці попарних порівнянь.....	51
3.6.3 Перевірка узгодженості оцінки експертів	55

3.7 Виконання зваженого накладання	56
3.8 Розробка моделі геообробки даних в середовищі ModelBuilder.....	58
4 СТВОРЕННЯ ІНТЕРАКТИВНОЇ КАРТИ	64
4.1 Веб-сервер GeoServer	65
4.2 База даних PostgreSQL/PostGIS.....	68
4.3 Бібліотека OpenLayers	68
ВИСНОВКИ.....	73
ПЕРЕЛІК ДЖЕРЕЛ ПОСИЛАНЬ.....	75
ДОДАТОК А Діаграма методології нечіткого зваженого накладання	78
ДОДАТОК Б Карти евклідових відстаней критеріїв.....	79
ДОДАТОК Г Модель геообробки даних	81

СКОРОЧЕННЯ ТА УМОВНІ ПОЗНАКИ

Скорочення

АРМ	– автоматизоване робоче місце
БГД	– база геоданих
ВММ	– віртуальна модель місцевості
ГІС	– географічна інформаційна система
ДДЗ	– дані дистанційного зондування
ІСАД	– інтелектуальні системи аналізу даних
МКАР	– мультикритеріальний аналіз рішень
ОПР	– особа, що приймає рішення
ПБГД	– персональна база геоданих
ПЗ	– програмне забезпечення
СКБД	– система керування базами даних
УВГ	– управління водним господарством
ЦМР	– цифрова модель рельєфу
DEM	– Digital Elevation Model (цифрова модель рельєфу)
ESRI	– Environmental Systems Research Institute
GRASS	– Geographic Resources Analysis Support System
GRS-80	– Geodetic Reference System, 1980
OGC	– Open Geospatial Consortium
OSGeo	– Open Source Geospatial Foundation
SAGA	– System for Automated Geoscientific Analyses
SDTS	– Spatial Data Transfer Standard
SRTM	– Shuttle Radar Topography Mission
uDig	– User-friendly Desktop Internet GIS
UTM	– Universal Transverse Mercator
WGS-84	– World Geodetic System, 1984

ВСТУП

Останнім часом великий інтерес науковців привертають методи, які дозволяють виконувати інтелектуальний аналіз даних (англ. Data mining), тобто виявляти в даних раніше невідомих, нетривіальних, практично корисних і доступних інтерпретацій знань, необхідних для прийняття рішень в різних сферах людської діяльності. У зв'язку з експоненціальним ростом кількості інформації, що потребує зберігання, обробки та аналізу, актуальним завданням сьогодення являється розробка моделей і методів, що забезпечують підтримку прийняття рішень (ППР), а також відповідного програмного забезпечення.

Велика частина проблем, що потребують ППР, мають просторовий характер. До них можна віднести такі проблеми як аналіз транспортної досяжності, аналіз небезпеки та ризику, аналіз придатності тощо. Протягом багатьох років дослідники та аналітики поєднують технології просторового аналізу та методів прийняття рішень, щоб допомогти вирішити такі проблеми. Запропоновані рішення дуже різноманітні. Деякі використовують інструменти послідовно, тоді як інші прагнуть запровадити спеціальні системи.

Геоінформаційні системи (ГІС) є ідеальною основою для інтелектуального аналізу даних, що мають просторову прив'язку, а також для розробки повнофункціональних систем підтримки просторових рішень. Це потужні системи управління, здатні збирати, аналізувати, моделювати та відображати просторові дані. Тому актуальним завданням є дослідження методів просторового моделювання з метою використання їх в інтелектуальному аналізі даних для підтримки прийняття рішень.

Метою даної магістерської роботи є дослідження і опис методів просторовому моделювання, що можуть використовуватися в інтелектуального аналізу даних, для вирішення завдання пошуку оптимального місця розташування об'єкту громадського харчування (кав'ярні) на території міста Одеса.

Для досягнення поставленої мети в роботі необхідно вирішити наступні завдання:

- виконати дослідження засобів ГІС для просторового моделювання;
- провести аналіз сучасних методів інтелектуального аналізу даних і підтримки прийняття рішень;
- обґрунтувати вибір методів, що будуть використовуватися для вирішення проблеми аналізу придатності;
- надати методологічні основи вирішення поставленої проблеми;
- розробити алгоритми та програмні реалізації етапів моделювання;
- створити інтерактивну карту проекту з метою сумісного редагування її декількома користувачами;
- виконати аналіз отриманих результатів.

Структура магістерської роботи складається з вступу, чотирьох розділів, висновків, переліку посилань на 28 найменувань, додатків. Повний обсяг роботи становить 81 сторінка, містить 17 рисунків і 3 таблиці.

1 ДОСЛІДЖЕННЯ ЗАСОБІВ ГІС ДЛЯ ПІДТРИМКИ ПРИЙНЯТТЯ ПРОСТОРОВИХ РІШЕНЬ

1.1 Географічні інформаційні системи

Геоінформаційні системи (ГІС) можуть розглядатися як системи, які дозволяють користувачу вводити, зберігати, керувати, аналізувати та представляти дані про просторові дані. Вони не обмежуються комп'ютерними інструментами та можуть включати ручні процедури.

Наведемо одне з безлічі визначень ГІС, що зустрічаються в літературі. Географічна інформаційна система являє собою комп'ютерну інформаційну систему, яка дозволяє здійснювати введення, моделювання, зберігання, вилучення, обмін, маніпулювання, аналіз та представлення географічно прив'язаних даних [1].

Для досягнення цих функцій ГІС об'єднує технологію бази даних, картографію та статистичний аналіз. Для введення даних в систему, функціональні можливості ГІС включають збор, переформатування, геоприв'язку та накопичення вказаних даних. Керування цими даними включає функції, які зазвичай доступні в системах керування базами даних (СКБД) для зберігання та вилучення інформації. Дійсно, більшість ГІС є орієнтованими на базу даних та мають можливість здійснювати інтегрований аналіз як просторових, так і атрибутних даних. Нарешті, ГІС дає змогу побачити дані у вигляді карт, таблиць, графіків, діаграм, тощо.

Існують протиріччя в судженнях різних авторів щодо функціональних можливостей ГІС, та належності її до конкретного типу систем. Деякі вважають, що ГІС є окремим випадком інформаційних систем або систем підтримки прийняття рішень, які були розширені для обробки географічних даних, тоді як інші вважають їх системами, розробленими з метою маніпулювання як просторовими, так і непросторовими даними. Оскільки ГІС використовується в безлічі дисциплін і предметів, це не дивно [1-3].

1.1.1 Моделі даних

ГІС використовують два типи даних: просторові дані та дані атрибутів [4]. Перший тип описує розташування просторових одиниць (наприклад, земельні ділянки, будинки, вулиці, річки, озера, муніципалітети, провінції, країни, тощо). Дані атрибутів відносяться до властивостей цих просторових сутностей і називаються табличними даними. Вони є кількісною та/або якісною інформацією, яку можна знайти в звичайних системах управління базами даних.

Просторові об'єкти можуть належати до однієї з трьох груп: точок, ліній та полігонів (або областей). Точки – це координати на карті, які можуть бути використані для представлення простих місць, таких як вершина, точки інтересу, тощо. З трьох типів просторових сутностей точки містять найменшу кількість інформації, оскільки неможливі вимірювання (тобто розмір або об'єм об'єкта). Точки, однак, можуть бути використані для представлення областей у меншому масштабі. Наприклад, міста на карті світу будуть представлені точками замість полігонів. Точки також можуть бути з'єднані для формування ліній, які використовуються для позначення лінійних об'єктів, таких як дороги, стежки, річки. У ліній вже є довжина. Нарешті, замкнуті лінії або двовимірні полігони використовуються для географічних об'єктів, які охоплюють певну поверхню, такі як озера, будівлі. Вони також можуть представляти межі різних областей, таких як ліси, міста, країни.

Геопросторові дані зазвичай діляться на два класи: векторні дані та растрові дані. Ці два типи моделей даних говорять про те, як дані зберігаються в системі.

Векторний формат для зберігання просторових даних використовує вектори, які є лінійними сегментами, визначеними їх кінцевими точками. Вектор розміру 0 представляє точку. Послідовність взаємопов'язаних векторів являє собою лінію, а ланцюжки векторів, які повертаються назад до їх відправної точки, можуть представляти межі області або полігону. На

відміну від растрової моделі, де простір явно дискретизується як сітка, векторна модель використовує координатування простору для визначення кінцевих точок кожного вектора. Використання векторних даних спрощує декілька операцій, коли існує необхідність об'єднати або масштабувати дані з різних джерел.

Растрові дані структуровані як масив або сітка комірок з пікселями. Растри можуть представляти будь-яку з трьох просторових одиниць, згаданих раніше (рис.1.1). Таким чином, точки можуть бути представлені одним пікселем, лінії – послідовністю з'єднаних пікселів, а області – групою сусідніх пікселів. Растри можуть полегшити обробку даних або систематичні операції, оскільки масиви зазвичай підтримуються мовами програмування. Проте вони неефективні з точки зору використовуваного об'єму пам'яті. Якщо одиниці, що представляються, не мають складних форм, відображення просторових об'єктів, що використовують растри, буде займати більше місця, ніж ті ж одиниці, що використовують векторну модель даних.

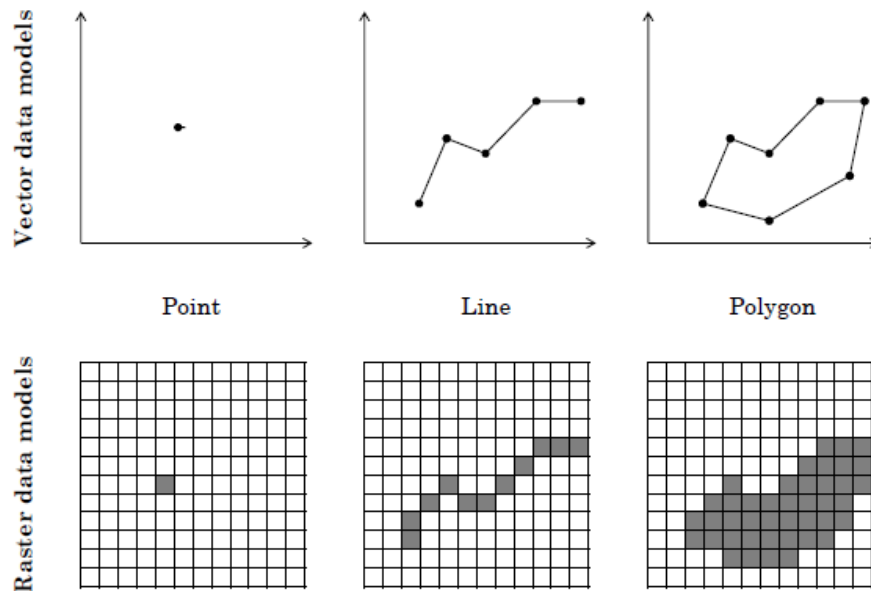


Рисунок 1.1 – Векторні та растрові моделі в ГІС

1.1.2 Приклади просторового аналізу

ГІС має велику панель інструментів просторового аналізу. Щоб продемонструвати кілька можливостей та функцій, що підтримуються ГІС, розглянемо ряд прикладів, які ілюструють їх через різні типи програм.

Інвентаризація ресурсів. За допомогою ГІС можлива інвентаризація просторово прив'язаних ресурсів. Як і будь-яка інша інформаційна система, ГІС здатна збирати дані з різних джерел, зв'язувати та поєднувати їх. За умови, що окремі джерела даних є сумісними за допомогою загальної просторової координатної системи, система зможе порівняти їх. Таким чином, ГІС може бути використана як інструмент для об'єднання даних з різних джерел інформації та управління ними.

Мережевий аналіз. Однією з найбільш затребуваних функцій ГІС є мережевий аналіз, який дозволяє системі управляти цілими мережами для транспортування, зв'язку і застосовувати на них різні алгоритми. При цьому система спроможна відповідати на запитання, наприклад, про найменший шлях до даного пункту призначення, обчисливши маршрут, який пов'язує декілька пунктів. Критеріями, що розглядаються тут, можуть бути відстань, але також час, який потрібно подорожувати, витрати на певні шляхи або пропускна здатність певних гілок в мережі за умови проблеми з потоком. В цьому аналізі використовуються алгоритми безпосередньо із теорії досліджень операцій, такі як методи знаходження найкоротшого шляху, задача комівояжера, або метод критичного шляху.

Розподілені дані. Подібно до збору даних з кількох джерел, ГІС може поширювати свої послуги в мережі зв'язку. Вхідні дані можна отримати з декількох веб-служб, користувачі можуть проглянути отриману інформацію на відстані, розрахункові маршрути можуть надсилатися на мобільні термінали, а результати просторового аналізу можуть бути розміщені на картах, які будуть надруковані. Таким чином, ГІС часто використовується, коли необхідна

спільна система, яка об'єднує декілька входів або залучає зацікавлених сторін до співпраці.

Територіальний аналіз. Територіальний аналіз базується на використанні даних рельєфу (рис. 1.2), щоб отримати інформацію про крутизну схилу. Ці моделі, які можна переглянути в трьох вимірах, можуть бути використані для виконання більш складних аналізів, таких як розрахунок водозбірних зон чи зон видимості. Водозбірні зони використовуються в гідрографії для прогнозування води в разі дощу або опадів. Зон видимості – це приклад аналізу видимості, який обчислює всі області, з яких видима точка. Вони можуть бути використані в дослідженнях в області регіонального планування для розташування вітрових турбін поза видимості жителів.

Аналіз на основі шару. На основі цього аналізу використовуються запити, які включають вибір за певними обмеженнями. Наприклад, знайти всі лікарні, що знаходяться в радіусі 5 км від аварії, або всі антени, що знаходяться в межах 0,5 км від основної дороги та всередині міста. Для проведення такого аналізу ГІС надає користувачу дві основні функції: буферизацію (*buffering*) та накладання (*overlay*). Буферизація означає утворення областей, які задовольняють обмеженням за відстанню, наприклад площа в межах 0,5 км від основної дороги. Буфери зазвичай круглі або прямокутні навколо точок і коридорів постійної ширини навколо ліній. На рисунку 1.3 показані приклади буферів для цих двох випадків. Логічне накладання стосується об'єднання, перетину, різниці або інших варіантів логічних операторів між декількома шарами. Це дозволяє ГІС поєднувати кілька обмежень для створення єдиної карти об'єктів, які їх задовольняють.

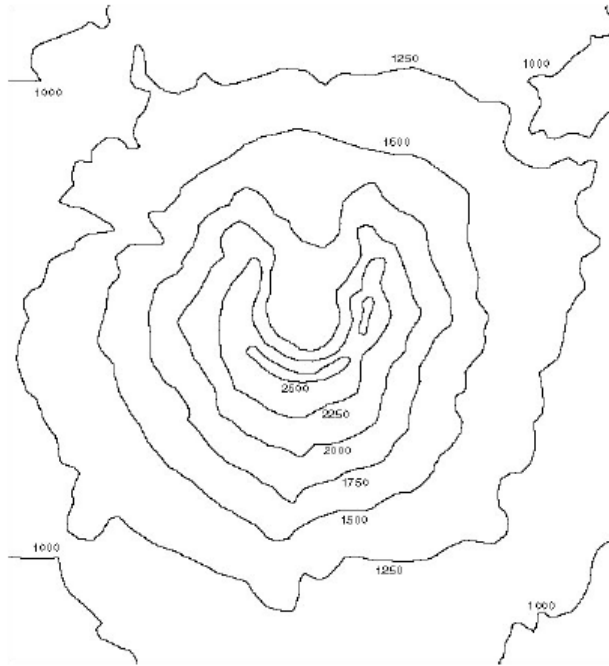


Рисунок 1.2 – Карта ізоліній, де кожна лінія має постійну висоту

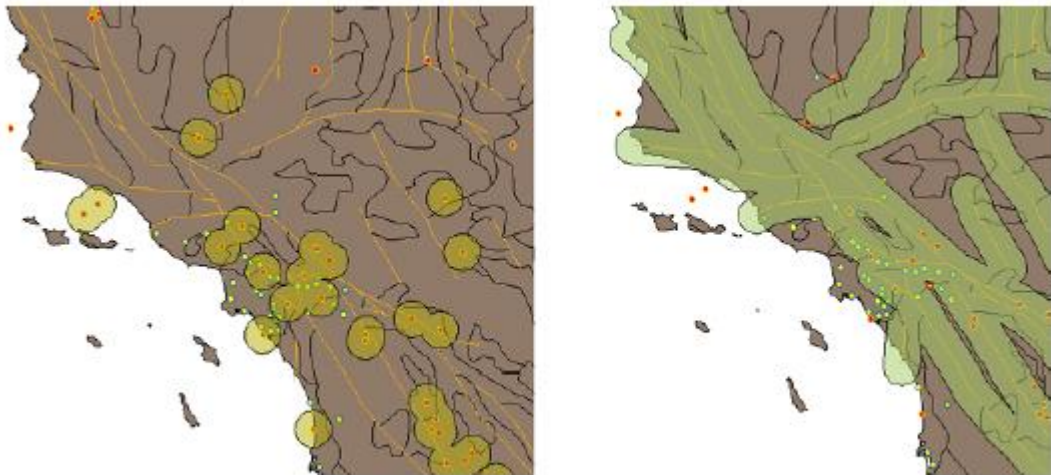


Рисунок 1.3 – Буферні карти для точок та ліній

Аналіз місця розташування. Моделі місцевості та аналіз шарів, описані вище, можуть бути використані для вибору місць розташування. Для їх виконання, ГІС пропонують інші функції, які служать для створення деяких шарів, які будуть використовуватися в аналізі. Одним із прикладів є створення

карт ізохронів, які показують сусідство певних місць, виходячи з відстані або часу, необхідного для їх досягнення. Інший приклад – використання проксимальних полігонів або діаграм Вороного/Тіссена. Вони відокремлюють місцевість у полігонах, які містять область, найближчу до певної точки. Приклад такого результату наведено на рис. 1.4.

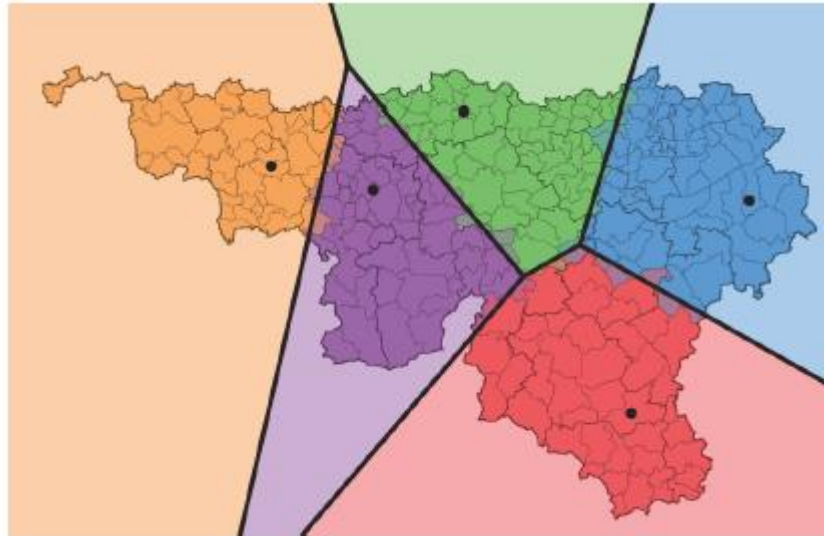


Рисунок 1.4 – Діаграма Вороного, яка отримана для п'яти точок

Просторово-часовий аналіз. Існує мало робіт щодо використання трьох видів вимірювань прив'язаних даних (наприклад, простору, часу та атрибуту). Незважаючи на те, що існуючі системи здатні збирати та зберігати часові дані, при аналізі та використанні їх для прогнозування, виникають декілька труднощів. Цей тип аналізу може привести до розуміння часової еволюції просторових явищ, якщо вони правильно оброблені.

На закінчення, в залежності від аналітичної обробки, яку вимагає аналіз, ми могли б розмістити їх в одній з трьох категорій:

1) Геометричний чи топологічний аналіз, якщо для аналізу потрібні логічні операції (наприклад, зв'язність, накладання) над просторовими об'єктами;

2) Аналіз атрибутів, якщо це пов'язано з варіаціями атрибутів області (наприклад, висота, тип ґрунту, швидкість вітру), які можуть бути дискретними або безперервними;

3) Мережевий аналіз, які включає в себе конфігурацію зв'язків між вузлами (наприклад, дорожньої мережі, мережами громадського транспорту, мережами дротових комунікацій, тощо).

1.2 Порівняльний аналіз сучасних ГІС пакетів

1.2.1 Програмний комплекс ArcGIS Desktop

Серед універсальних векторних ГІС особливе місце займає лінійка ГІС компанії ESRI (США).

Платформа ArcGIS 10 є оптимальним рішенням для побудови корпоративної ГІС. Вона побудована на основі стандартів комп'ютерної галузі, включаючи об'єктну архітектуру COM, .NET, Java, XML, SOAP, що забезпечує підтримку загальноприйнятих стандартів, гнучкість пропонуємих рішень, широкі можливості взаємодії. Фундаментальна архітектура ArcGIS забезпечує її використання у багатьох прикладних сферах і на різних рівнях організації роботи користувачів ГІС.

Система ArcGIS дозволяє виконувати всі основні функції по введенню, коригуванню, оновленню, зберіганню, обробки та подання просторових даних і побудована за модульним принципом. Кожна група винесена в окремий модуль розширення ArcGIS. У ArcGIS for Desktop є чотири основні додатки, які можна використовувати для картографування і візуалізації [5-7]:

1) ArcMap – основний додаток геоінформаційної системи ArcGIS для картографування, редагування, аналізу даних і управління ними. Він використовується для роботи з двовимірними об'єктами (рис.1.5).

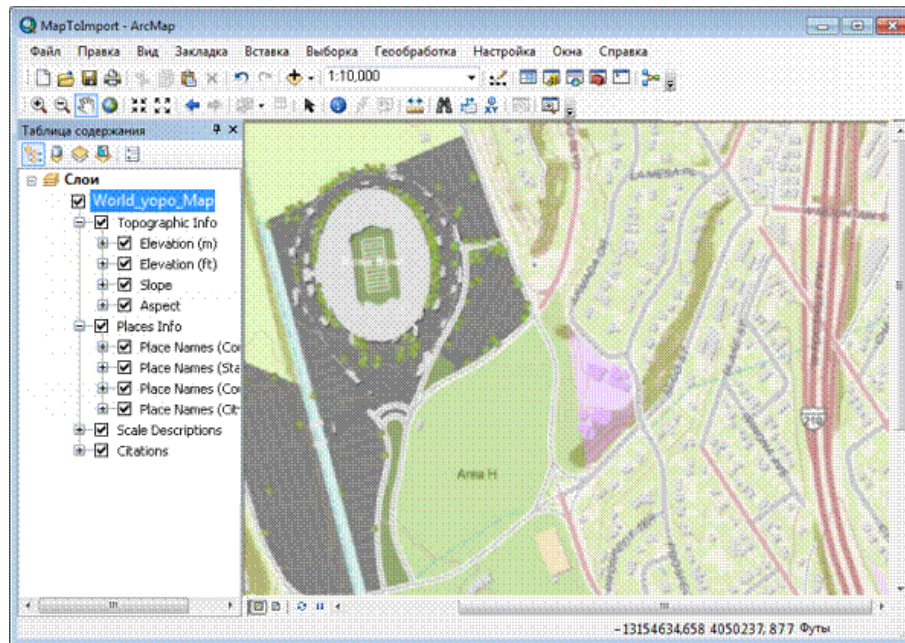


Рисунок 1.5 – Головне вікно програми ArcMap

2) ArcGlobe – додаток для візуалізації геоданих в тривимірному вигляді, для чого застосовується вид глобуса (рис.1.6). Глобально додаток спроектовано під роботу з дуже великими наборами даних, які повинні відображатися з різним рівнем детальності. ArcGlobe є опціональною частиною додаткового модуля ArcGIS 3D Analyst.

3) ArcScene – додаток для тривимірної візуалізації тривимірних сцен і ділянок. У ньому створюється тривимірний вигляд сцени, в якій можна виконувати різну навігацію (рис.1.7). ArcScene також є частиною додаткового модуля ArcGIS 3D Analyst.

4) ArcCatalog – це вікно каталогу, яке використовується для організації та управління різними типами географічної інформації для ArcGIS for Desktop. У ArcCatalog можна організувати і управляти наступними типами інформації: бази геоданих; файли растрів; документи карт, глобусів, 3D-сцен і файли шарів; набори інструментів геообробки, моделі, скрипти Python; ГІС-сервіси, опубліковані за допомогою ArcGIS for Server; метадані в певних стандартах для цих ГІС-елементів тощо. ArcCatalog представляє цю інформацію

у вигляді дерева і дозволяє вибрати ГІС-елемент, переглянути його властивості та отримати доступ до інструментів для роботи з ним (рис.1.8).

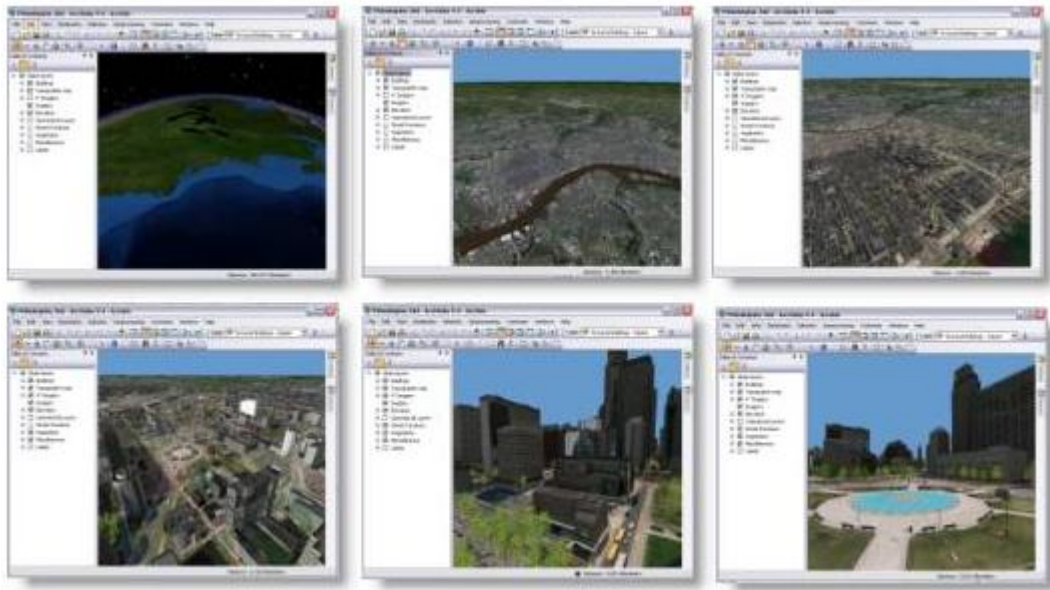


Рисунок 1.6 – Відображення інформації в ArcGlobe

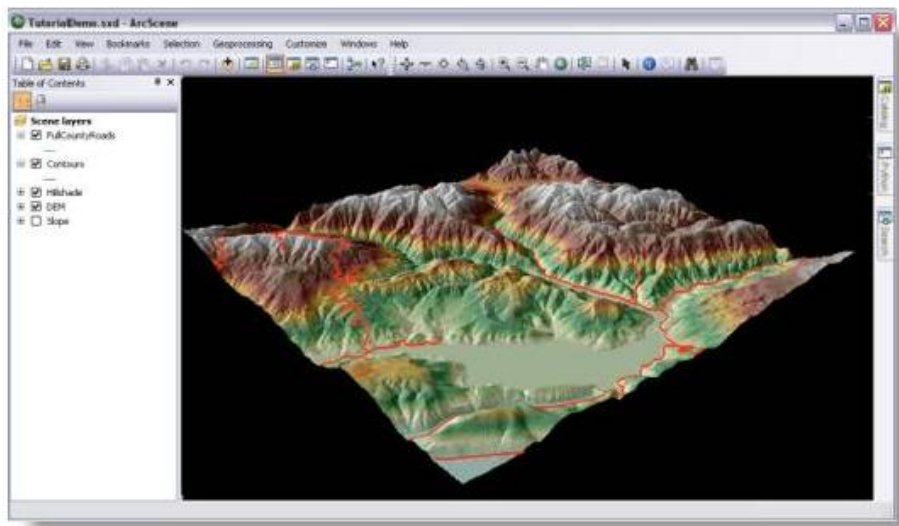


Рисунок 1.7 – Головне вікно програми ArcScene

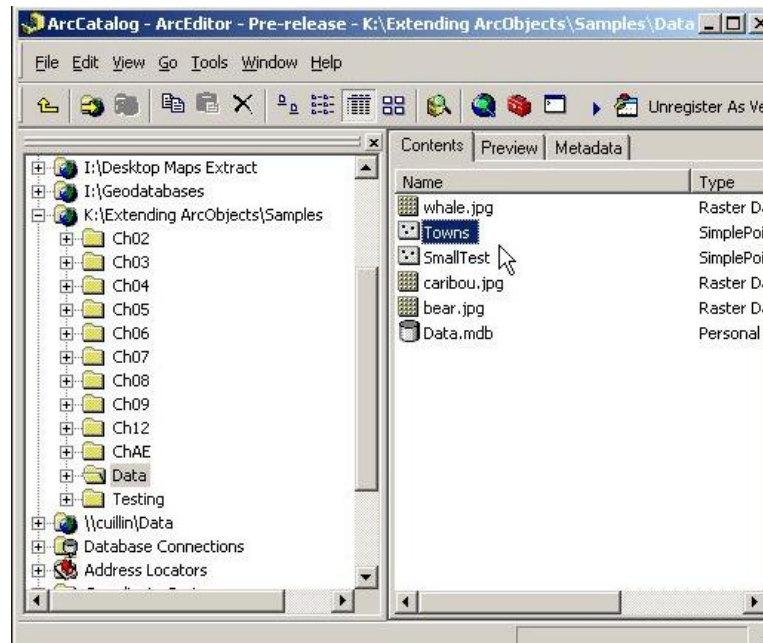


Рисунок 1.8 – Головне вікно програми ArcCatalog

1.2.2 Пакет ГІС MapInfo Professional

Універсальна ГІС MapInfo Professional розроблена компанією MapInfo Corp. (США) і орієнтована на широке коло користувачів. Система має великий набір функцій для роботи з не топологічними просторовими даними.

Для збереження даних можна використовувати СУБД ORACLE і MS SQL Server. ГІС MapInfo Professional має розвинуті засоби візуалізації і редагування картографічних даних (рис. 1.9), засоби тематичного картування (діапазони, діаграми, розмірні символи, щільність точок, індивідуальні значення).

Для роботи з растровими зображеннями в системі передбачено два режими: з реєстрацією і без неї. Перший використовується для сумісництва координатно-прив'язаних растрів з векторними даними, а другий – тільки для їх перегляду.

MapInfo Professional підтримує велику кількість графічних форматів, у тому числі тих, що використовуються для збереження даних ДЗЗ. У системі є

розвинуті засоби ділової графіки, що дозволяють формувати різні графіки (лінійні, стовбчасті, кругові, площадні, трьохвимірні). Всі графіки будуються за даними, що зберігаються у таблицях. Всі сформовані в ГІС данні (картографічні, табличні, ділова графіка і ін.) можуть бути виведені на друк. Для цього в системі є механізм створення макету друку.

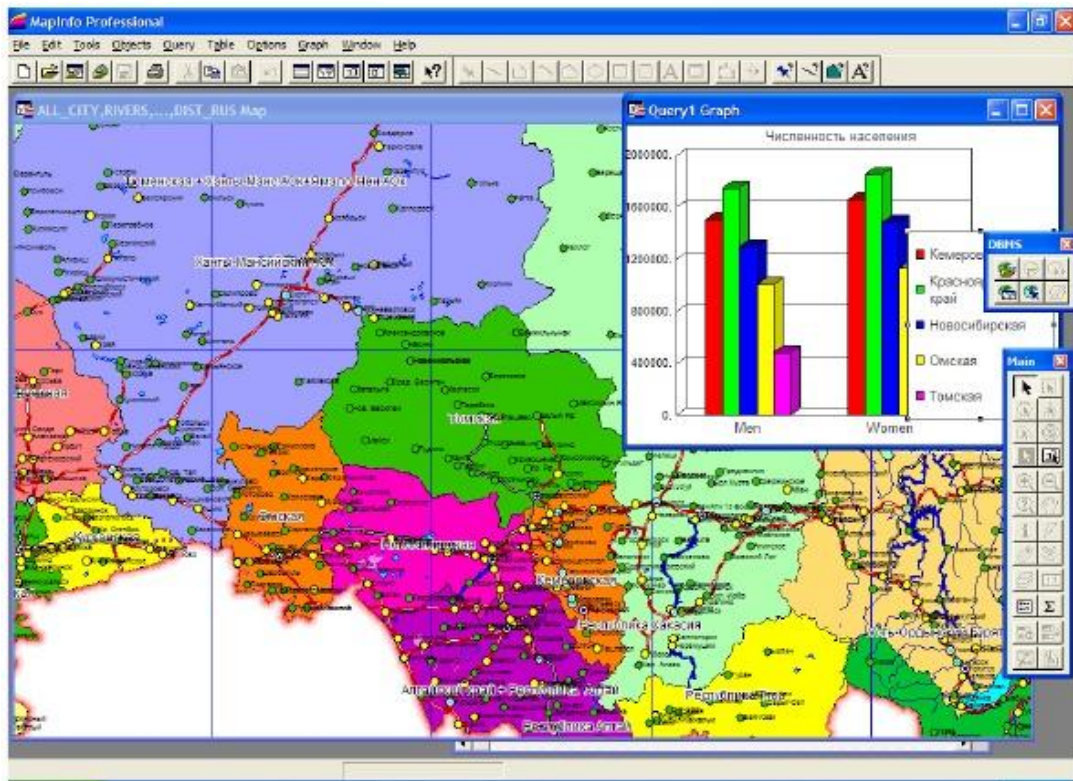


Рисунок 1.9 – Засоби візуалізації у ГІС MapInfo Professional

Для розширення можливостей цієї системи використовується спеціальна мова програмування MapBasic. За допомогою цієї мови можна модифікувати інтерфейс користувача, додавати до ГІС нові функції, керувати картами. У мові також є можливість підключення динамічних бібліотек (DLL). Компанія також пропонує додаткові модулі, що призначені для вирішення різних прикладних завдань.

1.2.3 Пакет QGIS і інші відкриті настільні ГІС

Визначення відкритого програмного забезпечення розроблено організацією Open Source Initiative і використовується для визначення відповідності ліцензії на програмне забезпечення стандартам відкритого ПЗ. Основні особливості відкритого ПЗ згідно з визначенням включають вільне поширення, доступний вихідний код, дозвіл на модифікацію цього вихідного коду [8].

Перелік FreeGIS.org [9] на даний момент налічує приблизно 356 відкритих програмних пакетів ГІС різного типу. Відкриті ГІС створюються і підтримуються різними спільнотами та організаціями: комерційними компаніями, групами ентузіастів або дослідницькими організаціями. Найбільші проекти залучають досить велику кількість розробників і вкладення в розробку досить серйозні (табл. 1.1).

Таблиця 1.1 – Характеристики програмної бази та оцінки витрат деяких відкритих ГІС

Проект	Рядків коду, тис. (приріст)	Розробників, чол.	Витрати, чол./рік	Оцінка витрат, тис. \$USD
GRASS GIS	737 (42%)	62	200	11000
gvSIG	2162 (20%)	62	609	33495
Q GIS	440 (227%)	40	114	6270
GDAL	1035 (67%)	29	337	18535

Справжній бум розвитку відкритих ГІС і пов'язане з ним поява нових користувальницьких систем припадає на початок ХХІ століття. У цей період почалася розробка таких користувальницьких ГІС, як SAGA GIS (2001) в Німеччині, gvSIG (2003) в Іспанії, міжнародний проект QGIS (2002). У 2007 пропріетарний раніше пакет, призначений для ГІС-аналізу та завдань дистанційного зондування ILWIS (Integrated Land and Water Information System)

офіційно став доступний під ліцензією GNU GPL, тим самим перейшовши до лав відкритого програмного забезпечення.

У 2006 році з метою підтримки і сприяння розробці проектів відкритих геопросторових технологій і баз даних з'явилася некомерційна організація Open Source Geospatial Foundation (скорочено OSGeo) [10]. Крім підтримки відкритих проектів, під егідою OSGeo випускається журнал, ведеться розробка і поширення навчальних матеріалів, проводяться щорічні міжнародні конференції (FOSS4G), присвячені відкритому ПЗ ГІС.

Ще одна організація, яка відіграє важливу роль у розвитку відкритих ГІС – Open Geospatial Consortium (OGC) [11]. Вона відповідальна за розробку стандартів взаємодії та обміну даними між різними ГІС-платформами. Крім університетів і адміністративних органів, членами OGC є в тому числі і розробники комерційних ГІС-платформ і баз даних. Наприклад стандарт, що описує інтерфейс доступу до географічних даних, що зберігаються в БД був реалізований як у відкритих базах даних (наприклад, PostgreSQL + PostGIS), так і в комерційних, таких як Oracle і IBM DB2.

Існуюче ПО ГІС можна умовно поділити на 3 класа: це веб ГІС, настільні ГІС і просторові бази даних. У табл. 1.2 представлені типові стеки відкритого ПЗ для настільних ГІС.

Розробка QGIS почалася в 2002 році групою ентузіастів. Метою розробки було створення простого у використанні і швидкого переглядача географічних даних для операційних систем сімейства Linux. Однак, із зростанням проекту з'явилася ідея використовувати QGIS як простий графічний інтерфейс для GRASS, отримуючи таким чином в своє розпорядження його аналітичні та інші функції. На сьогоднішній момент група розробників QGIS вирішила початкові завдання і працює над розширенням функціональних можливостей, давно вийшли за рамки простого переглядача. За рахунок використання крос-платформного тулкіта QT, QGIS доступна для більшості сучасних платформ (Windows, Mac OS X, Linux) і поєднує в собі підтримку векторних і растрових даних, а також здатна працювати з даними, що надаються різними

картографічними веб-серверами і багатьма поширеними просторовими базами даних. Функціональність QGIS може бути розвинена за допомогою створення модулів розширення на C++, або Python.

Таблиця 1.2 – Інструментальні шари відкритих настільних ГІС платформ

Тип ПЗ	Представники	Група
Додатки	QGIS, GRASS, OSSIM, uDig, MapWindow GIS	Користувальницький інтерфейс
Середовище розробки	Eclipse, QT, OpenGL, SharpDevelop	
Високорівневі утиліти	GeoTools, PostGIS, MapWindow GIS ActiveX	Збереження даних
Високорівневі скриптові мови програмування	Python, Perl, R	Обробка даних
Низькорівневі утиліти	Shapelib, JTS/GEOS, GDAL/OGR, GMT	
Низькорівневі мови програмування	C, C++, Java, Fortran, C#, VB.NET	Системне ПЗ
Операційна система	Linux, Microsoft Windows	

QGIS має одне з найбільш розвинених співтовариств в середовищі відкритих ГІС, при цьому кількість розробників постійно збільшується, чому сприяє наявність хорошої документації по процесу розробки і зручна архітектура.

2 АНАЛІЗ ПІДХОДІВ І МЕТОДІВ ВИРІШЕННЯ ЗАДАЧІ

2.1 Огляд методів Data Mining

Стрімкий розвиток інформаційних технологій, зокрема, прогрес в методах збору, зберігання і обробки даних дозволив багатьом організаціям збирати величезні масиви даних, які необхідно аналізувати. Обсяги цих даних настільки великі, що можливостей експертів вже не вистачає.

На сьогоднішній день інтенсивно розвивається напрямок, пов'язаний з інтелектуалізацією методів обробки та аналізу даних. Інтелектуальні системи аналізу даних (ІСАД) покликані мінімізувати зусилля особи, яка приймає рішення (ОПР), в процесі аналізу даних, а також в налаштуванні алгоритмів аналізу. Багато ІСАД дозволяють не тільки вирішувати класичні задачі прийняття рішення, а й здатні виявляти причинно-наслідкові зв'язки, приховані закономірності в системі, яку піддають аналізу.

Основу методів Data Mining складають різні методи класифікації, моделювання і прогнозування. До методів Data Mining нерідко відносять статистичні методи (дескриптивний аналіз, кореляційний і регресійний аналіз, факторний аналіз, дисперсійний аналіз, компонентний аналіз, дискримінантний аналіз, аналіз часових рядів). Такі методи, проте, припускають деякі апіорні уявлення про аналізовані дані, тому виникає певна розбіжність з цілями Data Mining (виявлення раніше невідомих нетривіальних і практично корисних знань).

Одне з найважливіших призначень методів Data Mining полягає в наочному поданні результатів обчислень, що дозволяє використовувати інструментарій Data Mining людьми, які не мають спеціальної математичної підготовки. У той же час, застосування статистичних методів аналізу даних вимагає доброго володіння теорією ймовірностей і математичної статистики.

Знання, що видобуваються методами Data mining, прийнято представляти у вигляді наступних моделей:

- асоціативні правила;
- дерева рішень;
- кластери;
- математичні функції.

Методи побудови таких моделей прийнято відносити до області штучного інтелекту. До методів і алгоритмів Data Mining відносяться:

- штучні нейронні мережі;
- дерева рішень, символні правила;
- методи найближчого сусіда і k-найближчого сусіда;
- метод опорних векторів;
- байєсовські мережі;
- лінійна регресія;
- кореляційно-регресійний аналіз;
- ієрархічні методи кластерного аналізу;
- неієрархічні методи кластерного аналізу, в тому числі алгоритми k-середніх і k-медіани;
- методи пошуку асоціативних правил, в тому числі алгоритм Apriori;
- метод обмеженого перебору;
- еволюційне програмування і генетичні алгоритми;
- різноманітні методи візуалізації даних і безліч інших методів.

Окремо слід зупинитися на логічних методах, або методах логічної індукції, які включають: нечіткі запити і аналізи; символні правила; дерева рішень; генетичні алгоритми. Методи цієї групи є, мабуть, таким, що найбільш інтерпретуються – вони оформляють знайдені закономірності, в більшості випадків, в досить прозорому вигляді з точки зору користувача. Отримані правила можуть включати безперервні і дискретні змінні. Слід зауважити, що дерева рішень можуть бути легко перетворені в набори символних правил шляхом генерації одного правила по шляху від кореня дерева до його термінальній вершини. Дерева рішень і правила фактично є різними спосо-

бами вирішення однієї задачі і відрізняються лише за своїми можливостями. Крім того, реалізація правил здійснюється більш повільними алгоритмами, ніж індукція дерев рішень.

2.2 Огляд методів мультикритеріального вибору

Вибір та прийняття рішень, часто, відбуваються в умовах мультикритеріальності, нечіткості та неповноти вихідної інформації. Оскільки рішення так чи інакше повинно бути прийнято, то недостача інформації може бути заповнена лише експертами на підставі їх досвіду та інтуїції. При недостачі об'єктивної кількісної інформації, ефективність будь-якого метода мультикритеріального вибору залежить від засобу представлення і використання знань експертів. Тому, необхідні підходи, що враховують суб'єктивні оцінки, знання та переваги експертів.

Для моделювання знань та досвіду експертів були розроблені різні як чіткі, так і нечіткі моделі представлення знань, моделі та методи підтримки рішень, серед яких найбільшій інтерес представляють нечіткі ситуаційні моделі, спеціальні семантичні і фреймові мережі, класифікаційні та композиційні моделі, моделі, що використовують нечіткі відношення переваги, якісне ранжування альтернатив, аналіз когнітивних карт і аналіз ієрархій, та ін. [12-14].

Найбільш поширеними нечіткими моделями прийняття рішень є моделі, що використовують нечіткі відношення та композицію нечітких функцій, а також класифікаційні моделі. Композиційні моделі використовуються як при формалізації наближених суджень ОПР, так і при організації логічного висновку в експертних системах, де знання представлені у вигляді набору нечітких продукцій [14].

В останні десятиріччя швидко розвивається науковий напрямок, що досліджує особливості застосування методів мультикритеріального аналізу рішень (МКАР) в ГІС для підтримки прийняття рішень в задачах, що мають

просторових характер. Далі розглянемо окремі методи, що можуть бути реалізовані в географічному контексті.

2.2.1 Зважена сума (Weighted Sum)

Метод зваженої суми (Weighted Sum) є найбільш популярним методом мультикритеріального рішення завдяки його простоті. Як зазначено в назві, це просто сума зважених оцінок:

$$p_a = \sum_{i=1}^n x_{ai} w_i \quad (2.1)$$

де p_a – пріоритетна оцінка альтернативи a ,

x_{ai} – оцінка альтернативи a за критерієм i ,

w_i – вага критерію i , n – кількість критеріїв, припускаємо, що метою є максимізація усіх критеріїв.

Завдяки своїй простій і дуже зрозумілій методології, зважене сума підходить для вирішення таких проблем, в яких особа, яка приймає рішення, більшою мірою бере участь у етапі моделювання. Дійсно, підхід передбачає можливість візуалізації проміжних результатів, які можуть допомогти підтримати та покращити переговори між зацікавленими сторонами. Багато програмних засобів реалізують такий підхід і пропанують додаткові інструменти, такі як аналіз чутливості та можливості візуалізації [15, 16].

Однак зважена сума має деякі недоліки [17]. Одне з них – це кілька припущень, які потрібно зробити перед використанням:

- Компенсація між критеріями: всі критерії знижуються і виражаються в одних і тих же одиницях (в цьому випадку через функції вартості). Це означає, що погана продуктивність за критерієм (наприклад, високий вплив на навколишнє середовище) може бути компенсована хорошою продуктивністю за іншим (наприклад, високий дохід).

– Адитивність атрибутів: зважене сумування може застосовуватися тільки в тому випадку, якщо атрибути є адитивними. Це припущення означає, що не повинно бути ніякої взаємодії між атрибутами, тобто атрибути повинні бути незалежними один від одного, що в багатьох випадках є нереалістичним припущенням.

Застосування моделі зваженої суми також призводить до деяких розбіжностей, які потрібно вирішити.

- втрата інформації внаслідок нормалізації;
- призначення ваги – складне завдання, особливо якщо кількість критеріїв велике і критерії дуже різні за своїм характером. Для підтримки осіб, які приймають рішення, були розроблені різні методи оцінки відносної важливості критеріїв.

2.2.2 Метод аналізу ієрархій (АНР)

Метод аналізу ієрархій (АНР), розроблений Томасом Л. Саати, є відомим інструментом для надання мультикритеріальної допомоги щодо прийняття рішень [18-20]. Він широко використовується в різних галузях: машинобудування, виробництва, промисловості, логістики, охорони здоров'я та ін. Однією з відмінних особливостей АНР є створення матриці парних порівнянь критеріїв з використанням вербальної шкали. У стандартній версії методу нормалізований власний вектор цієї матриці дозволяє обчислювати оцінку, пов'язану з кожною альтернативою, і вагу, пов'язану з кожним критерієм. Далі розглянемо, як АНР може бути використаний для встановлення розподілу ваги за критеріями. Аналогічний підхід використовується для визначення оцінок альтернатив. Ця методика широко використовується в задачах просторового аналізу і навіть існує в деяких комерційних ГІС або як скрипт, або як модуль МКАР.

Одним з найважливіших кроків методу АНР є побудова матриці A , де кожен елемент A_{ij} ($i, j = 1, \dots, n$) представляє відносну важливість критерію i в

порівнянні з критерієм j . Для того, щоб висловити цю відносну важливість, особа, що приймає рішення, може скористатися словесним масштабом. Останній потім перетворюється в фундаментальну шкалу абсолютних чисел, що приймає цілі значення від 1 до 9. Повне пояснення цієї шкали можна знайти в [20]. Крім того, елементи цієї матриці повинні дотримуватися наступної властивості узгодженості:

$$A_{ij} = \frac{1}{A_{ji}}, \quad \forall i, j \quad (2.2)$$

При використанні 9-бальної шкали, матриця A приймає свої значення в наступному діапазоні $\left\{\frac{1}{2}, \frac{1}{3}, \dots, \frac{1}{9}, 1, 2, \dots, 9\right\}$.

В ідеальному випадку матриця є послідовною, тобто вона, природно, має таку властивість:

$$A_{ij} = A_{ik} \times A_{kj} \quad \forall i, j, k \quad (2.3)$$

У цьому випадку матрицю попарних порівнянь A можна записати наступним чином:

$$A = \begin{bmatrix} w_1/w_1 & w_1/w_2 & \cdots & w_1/w_n \\ w_2/w_1 & w_2/w_2 & \cdots & w_2/w_n \\ \vdots & \vdots & \ddots & \vdots \\ w_n/w_1 & w_n/w_2 & \cdots & w_n/w_n \end{bmatrix} \quad (2.4)$$

де w_1, w_2, \dots, w_n є вагами критеріїв.

Щоб знайти вектор $w(w_1, w_2, \dots, w_n)$, можна вирішити наступне рівняння:

$$A_w = n w \quad (2.5)$$

де вектор w – власний вектор A .

Для контролю узгодженості експертних оцінок вводяться дві пов'язані характеристики – індекс узгодженості (Consistency Index, C.I.) і відношення узгодженості (Consistency Ratio, C.R.) [19]:

$$C.I. = \frac{\lambda_{max} - n}{n-1},$$

(2.6)

де λ_{max} – найбільше власне значення.

Томас Саати також встановив набір контрольних значень R.I., середніх відношень узгодженості, які залежать від кількості альтернатив [19]. Коли C.I. менша або рівна 10% від R.I., рівень узгодженості вважається прийнятним для застосування методу власного вектора.

$$C.R. = \frac{C.I.}{R.I.},$$

(2.7)

де R.I. – середнє відношення узгодженості, яке залежить від розміру вибірки. Розумний рівень узгодженості в парних порівняннях передбачається на рівні C.R. < 0.10, якщо C.R. \geq 0.10, це показує суперечливі судження.

Процес встановлення матриці оцінки такий самий, як і розглянутий вище. Різниця полягає лише в тому, що замість порівняння критеріїв на підставі їх важливості, особам, що приймають рішення, буде запропоновано порівняти альтернативи на основі їхніх уподобань.

2.2.3 Нечітка логіка

При роботі з растровими картами часто використовується булава логіка.

Булава модель містить логічну комбінацію бінарних відображень, що виникає в результаті застосування умовних операторів. Кожній комірці растру присвоюється значення «1» або «0», вказуючи, чи є вона задовільною або незадовільною, відповідно до цілі моделювання.

На відміну від класичної двійкової логіки, в якій певний елемент є істинним або хибним, а об'єкт належить множині, чи ні, нечітка логіка, що була введена Лютфи Заде в 1965 році [21] допускає існування функції належності елемента до множини, яка приймає значення у інтервалі $[0, 1]$.

Введемо визначення нечіткої множини згідно Л. Заде [21].

Нехай U – універсальна множина, u – елемент U . Нечітка множина \tilde{A} універсальної множини U визначається як множина впорядкованих пар

$$\tilde{A} = \{\{\mu_A(u)|u\}\}, \quad (2.8)$$

де $\mu_A(u)$ – функція належності, що приймає значення в деякій цілком впорядкованій множині $M = [0; 1]$. Якщо $M = \{0, 1\}$, то нечітка підмножина може розглядатися як чітка множина.

Функція належності вказує ступінь належності елемента u нечіткій підмножині \tilde{A} . Чим більше $\mu_A(u)$, тим більшою мірою елемент універсальної множини відповідає властивостям нечіткої підмножини.

В теорії нечітких множин функція належності відіграє значну роль, так як це основна характеристика нечіткого об'єкта, а всі дії з нечіткими об'єктами виконуються через операції з їх функціями належності. Як правило, функція належності будується за участю експерта (групи експертів), таким чином, що ступінь належності приблизно дорівнює інтенсивності прояву деякої властивості.

Для опису об'єктів в умовах невизначеності використовується поняття нечіткої змінної, яка задається трійкою

$$\langle \alpha, U, \tilde{A} \rangle,$$

де α – найменування нечіткої змінної, $U = \{u\}$ – область визначення змінної α , $\tilde{A} = \{\{\mu_A(u)|u\}\}$ – нечітка множина, що описує обмеження на можливості значення змінної α . Нечітка змінна – це теж саме, що і нечітке число, тільки з додаванням імені, яким формалізується поняття, що описується цим числом.

Лінгвістична змінна – це множина нечітких змінних. Вона використовується для того, щоб дати словесний опис деякому нечіткому числу, отриманому в результаті деяких операцій і задається кортежем

$$\langle \beta, U, T, M, G \rangle,$$

де β – назва лінгвістичної змінної, U – універсальна множина, T – множина значень лінгвістичної змінної, кожне з яких є нечіткою змінною на U , M – синтаксична процедура, що дозволяє отримувати (формувати) нові значення лінгвістичної змінної, G – семантичне (сміслове) правило, що дозволяє інтерпретувати знову одержувані значення лінгвістичної змінної.

Терм-множина – це множина всіх можливих значень лінгвістичної змінної.

Терм – будь-який елемент терм-множини. У теорії нечітких множин терм формалізується нечіткою множиною за допомогою функцій належності.

В теорії нечітких множин функція належності відіграє значну роль, так як це основна характеристика нечіткого об'єкта, а всі дії з нечіткими об'єктами виконуються через операції з їх функціями належності. Як правило, функція належності будується за участю експерта (групи експертів), таким чином, що ступінь належності приблизно дорівнює інтенсивності прояву деякої властивості.

На практиці застосовуються такі види функцій належності [22]: лінійні, трикутні і трапецеїдальні (лінійно-кускові); нелінійні (функція Гауса, сигмоїдальна функція, сплайн).

2.2.4 Метод зваженого лінійного комбінування

Модель зваженої лінійної комбінації (WLC) є одним з найбільш широко використовуваних правил прийняття рішень на основі ГІС [23]. Цей метод часто застосовується при аналізі землекористування/придатності, вибору ділянки і проблеми оцінки ресурсів. Основною причиною його популярності є те, що цей метод легко реалізувати в середовищі ГІС з використанням операцій картографічної алгебри і картографічного моделювання. Цей метод також легко зрозумілий і інтуїтивно привабливий для осіб, які приймають рішення.

WLC можна формалізувати за допомогою мультикритеріального прийняття рішення. Нехай набір альтернатив рішення буде представлений:

$$X = \{x_{i*} | i = 1, 2, \dots, m\} \quad (2.9)$$

Альтернативи представлені набором комірок або пікселів в растровій базі даних ГІС. Таким чином, індекс i вказує місце розташування i -ої альтернативи. Для простоти будемо використовувати один індекс для вказівки місця розташування альтернативи (комірка, що позначена $i = 1$, знаходиться в верхньому лівому куті сітки карти, комірки пронумеровані зліва направо, справа для кожного рядка, комірка m розташована в нижньому правому куті растрової карти). Кожна альтернатива описується за допомогою її локального атрибута (дані координат) і атрибута даних (значення атрибутів, пов'язані з місцем розташування). Оскільки атрибути служать змінними рішення, можемо призначити результат рішення (значення критерію) як x_{ij} , який представ-

ляє рівень j -го атрибута щодо альтернативи i . Отже, альтернативу i можна охарактеризувати вектором в рівнянні (2.10), а рівні атрибутів по альтернативі представлені вектором в рівнянні (2.11).

$$x_{i*} = (x_{i1}, x_{i2}, \dots, x_{in}), \quad i = 1, 2, \dots, m \quad (2.10)$$

$$x_{*j} = (x_{1j}, x_{2j}, \dots, x_{mj}), \quad j = 1, 2, \dots, n \quad (2.11)$$

Вхідні дані для рівнянь (2.10) і (2.11) можуть бути організовані у вигляді таблиць (оціночна матриця або географічна матриця). Відповідно, дані можуть зберігатися в ГІС як набір шарів карти. Дані складаються з набору n шарів даних, і кожна комірка сітки в шарі даних містить значення атрибута x_{ij} . У конкретній ситуації прийняття рішення набір альтернатив може бути обмежений шляхом накладення обмежень на значення атрибутів (непросторове обмеження) або на атрибути розташування (просторові обмеження).

З огляду на вхідні дані, проблема полягає в об'єднанні (агрегації) шарів карти відповідно до правила вирішення WLC. Формально правило прийняття рішення оцінює кожну альтернативу, наступною функцією цінності:

$$V(\mathbf{x}_i) = \sum_j w_j v_j(\mathbf{x}_i) = \sum_j w_j r_{ij} \quad (2.12)$$

де w_j – нормалізована вага, така, що $\sum w_j = 1$;

$v_j(x_i)$ є функцією цінності для j -го атрибута, $x_i = (x_{i1}, x_{i2}, \dots, x_{in})$;

r_{ij} – атрибут, перетворений до порівняного масштабу.

Ваги представляють відносну важливість атрибутів. Найбільш краща альтернатива вибирається шляхом визначення максимального значення $V(\mathbf{x}_i)$ для $i = 1, 2, \dots, m$.

WLC-модель може бути модифікована для застосування ГІС шляхом включення карт обмежень в рівняння (2.12):

$$V(\mathbf{x}_i) = (\sum_j w_j r_{ij})(\prod_j r_{ik}^*), \quad (2.13)$$

де r_{ik}^* – значення, присвоєне i -й комірці на k -му шарі карти обмежень, воно приймає значення 1, якщо комірка є допустимою альтернативою, а значення 0 присвоюється нездійсненим коміркам.

З урахуванням правила прийняття рішення (2.13) метод WLC включає в себе наступні етапи:

- 1) Визначити набору атрибутів (цілей і пов'язаних шарів атрибутів карти).
- 2) Визначити набору можливих альтернатив.
- 3) Побудова співмірних карт атрибутів, їх стандартизація.
- 4) Визначення ваги критеріїв (тобто ваги «відносної важливості» безпосередньо прив'язаної до кожного атрибуту).
- 5) Об'єднання співмірних карт атрибутів і ваги за допомогою операцій накладення множення і сумування для отримання загальної оцінки для кожної комірки (альтернативи).
- 6) Ранжування альтернативи відповідно до загальних показників ефективності (комірка з найвищим балом є «найкращою» коміркою).

WLC може бути введений в дію з використанням будь-якої системи ГІС, що має можливості накладення. Методи накладення дозволяють об'єднати шари карти атрибутів (вхідні карти), щоб визначити складений шар карти (вихідну карту).

Щоб вирішити, чи підходить певна територія для певного об'єкта чи діяльності, необхідно визначити критерії, які допоможуть особам, що приймають рішення. Ці критерії (атрибути), які представлені у вигляді тематичних шарів ГІС, повинні відповідати певним вимогам. Відповідно [23], критерії повинні бути всебічними, вимірюваними, повними, не зайвими і мінімальними. Комплексний атрибут – це атрибут, який вказує на досягнення мети, по-

в'язаної з вирішенням проблеми. Вимірюваний атрибут – це атрибут, який можна оцінити, використовуючи номер, який вимірює переваги у вирішенні проблеми. Атрибути вважаються повними, якщо вони охоплюють всі аспекти проблеми. Не-надлишкові атрибути – це атрибути, незалежні один від одного таким чином, що кожен атрибут представляє один аспект проблеми. Вибір критеріїв базується на різних факторах, таких як ціль дослідження, експертний аналіз та наявність даних.

3 ПОБУДОВА МОДЕЛІ

3.1 Постановка завдання

Рішення про місце розташування об'єкта є методом багатокритеріального рішення, який широко використовується як приватними, так і громадськими організаціями. Це процес прийняття рішень органами, що оцінюють можливі площі, для визначення перспективних місць для різних видів об'єктів, таких як торговий центр, склад, лікарня, майданчик для зберігання та готель.

У секції роздрібно́ї торгівлі рішення про місцезнаходження є основною частиною для комерційної діяльності, оскільки це не тільки впливає на поточну дохідність, але також є ключовим етапом для подальшої розробки стратегії, яка безпосередньо впливає на довгострокові результати. Бізнес-модель може бути змінена шляхом перетворення зовнішнього середовища, однак, як тільки місце знайдено, це буде зробити дуже важко. Для малого бізнесу рішення про місце розташування розглядається як майже вирішальний компонент для його успіху або невдачі. У багатьох дослідженнях було показано, що проблеми вибору місця, складно вирішити традиційними простими методами із статичними моделями, наприклад, за попереднім досвідом, попередніми розрахунками та перевагами. Через відсутність динамічного аналізу складних та абстрактних факторів місцевості такі підходи не можуть вважатися науковими достатніми для прийняття обґрунтованих стратегічних рішень. З розвитком інформаційних технологій була запропонована і розроблена більш систематизована методологія, методика аналізу рішень на базі ГІС (GIS–MCDA).

ГІС-технології дозволяють обробляти та аналізувати велику кількість складних просторових даних, які не можуть бути використані в таблицях та формах, що призводить до їх специфічної ролі в оптимізації інформації для спрощення прийняття рішення. Залежно від здатності потужного графічного

представлення, ефективної організації даних та просторового аналізу, технологія на базі ГІС відіграє важливу роль у процесі вибору місця розташування комерційних об'єктів і стала важливою платформою візуалізації для аналізу проблем локалізації. Функція візуалізації ГІС, яка спрощує складні дані та усвідомлює взаємодію між альтернативами та відповідальними особами, чудово підходить для прийняття рішення.

Хороший вибір місця для бізнесу часто включає в себе кілька додаткових критеріїв, таких як розподіл ресурсів, ринковий потенціал, транспортна розв'язка, топографічні характеристики, вплив на навколишнє середовище тощо. Деякі з цих критеріїв несумірні і не можуть бути легко об'єднані в загальну одиницю. Крім того, традиційні способи прийняття рішень рідко включаються в різні способи мислення та нематеріальні та суперечливі чинники. Все це свідчить про те, що значні дослідження мультикритеріального аналізу рішень, забезпечують можливість прийняття кращого рішення серед класифікованих альтернатив за складними ознаками. Він "надає багатий набір методів та процедур для структурування проблем, що вирішуються, а також розробляє, оцінює та визначає пріоритети альтернативних рішень" [23].

Метою даної роботи є пошук оптимальних місць розташування для кав'ярень у місті Одеса, з використанням ГІС і мультикритеріальної моделі прийняття рішення, заснованої на методі аналізу ієрархій.

Для досягнення поставленої мети необхідно на першому кроці узагальнити необхідні фактори, які в основному впливають на місце розташування кафе.

3.2 Вихідні дані моделювання

Через наявність даних і те, що отримані результати повинні бути можливі для використання в інших областях, місто Одеса було обрано в якості тестової зони.

Одеса розташована на північно-західному узбережжі Чорного моря (рис.3.1). Площа міста сягає 162,42 км², густина населення — 6223 особи/км². За чисельністю населення Одеса є третім у країні містом після Києва та Харкова (станом на 1 травня 2017 року в Одесі проживало 1 010 912 осіб[24]). Одеса – культурний, освітньо-науковий, туристичний та торговий осередок країни; найбільший морський торговий порт країни; великий автомобільний та залізничний вузол. В Одесі діють: машинобудівна, хімічна, нафтопереробна, харчова та легка промисловості. Також налічується 37 вищих навчальних закладів, шість театрів, вісім кінотеатрів, філармонія, цирк, музеї, галереї тощо.

У місті велика кількість ресторанів. Станом на 2010 рік в Одесі наявні 648 об'єктів ресторанного бізнесу: ресторанів, кав'ярень, барів тощо. У місті розвинена інфраструктура роздрібної торгівлі. Мережа роздрібної торгівлі становить 14 670 об'єктів.

Для моделювання були використані векторні шари карти, отриманої з картографічного сервісу OpenStreetMap (OSM) [25].



Рисунок 3.1 – Карта міста Одеса

В роботі для обробки даних використовувалися програмні продукти: QGIS для отримання даних з OSM і ArcGIS для перетворення векторних даних в растрові дані, а також проведення просторового моделювання.

Оскільки базова карта була завантажена безпосередньо з OSM, у кожного об'єкту міститься багато атрибутів, і деякі з них можуть бути не обов'язковими. Фактори, які впливають на місце розташування кав'ярні, були обрані для створення нових шарів критеріїв. А потім перетворені з векторного на растровий формат, щоб їх можна було змінювати відповідно моделі.

3.3 Визначення критеріїв

У емпіричних дослідженнях, що було проаналізовано, при виборі місця розташування ресторану чи кав'ярні, завжди враховують такі чинники, як кількість трафіку, доступність, видимість, розмір кав'ярні та демографічні показники. Важлива транспортна розв'язка та наявність парковки. Для конкурентоспроможності ділянки важлива також наявність прямих чи непрямих конкурентів. В наукових публікаціях важливими факторами для аналізу місць розташування ресторанів називають характеристики району, комунальні послуги, ринок, конкуренцію, видимість, інформацію про дорожній рух, доступність послуг, вартість і тип обслуговування ресторану.

У цьому дослідженні процеси вибору місця були здійснені за географічними ознаками, внаслідок обмеження доступності даних основні розглянуті фактори належать до видів транспортної доступності, доступу покупців та конкуренції (додаток А).

До критеріїв транспортної доступності були віднесені пішохідні дороги, автобусні зупинки, головні дороги. Далі було зроблено припущення, що найбільша кількість покупців – це люди молодого віку, та ті, що знаходяться на відпочинку, тому до критеріїв доступу покупців були віднесені відстані від навчальних закладів, парків та пляжних зон. До обмежуючого фактору

відносяться існуючі кав'ярні, що можуть створювати конкуренцію, та наявність будівлі, тому що не передбачається будівництва нової.

3.4 Побудова обмежуючого шару

В якості обмежуючого шару був використаний шар з зображенням будівель міста Одеса. Для цього за SQL запитом до бази даних були вилучені до окремого шару об'єкти зі значенням атрибуту «building». Отриманий векторний шар об'єктів наведений на рис.3.2.



Рисунок 3.2 – Векторний шар з будівлями м. Одеси

Окремо були додані в обмежуючий шар будівлі у яких вже розміщуються діючі кав'ярні, ресторани, бари та точки фаст фуду (рис.3.3).

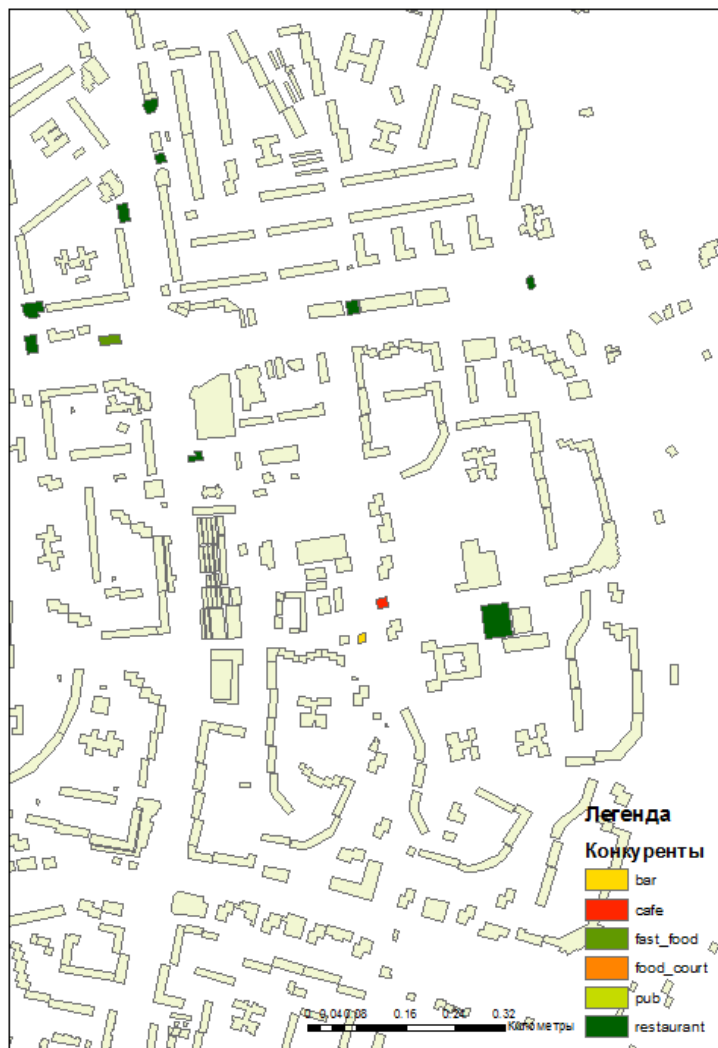


Рисунок 3.3 – Обмежуючий шар зі закладами конкурентами

3.5 Побудова шарів критеріїв та їх стандартизація

Для створення шару, який дасть змогу оцінити кількісно наявність потенційних покупців було прийнято рішення об'єднати в один шар полігональні об'єкти, що відображають школи, коледжі і вищі навчальні заклади м. Одеси, а також за допомогою SQL запиту до бази даних створити шари з парковими і пляжними зонами. Об'єкти отриманих шарів представлені на рис. 3.4.

За припущенням, що найбільша кількість покупців може бути в зоні найменшої відстані від зазначених об'єктів, були знайдені евклідові відстані від кожного типу об'єктів окремо (рис. Б.1).

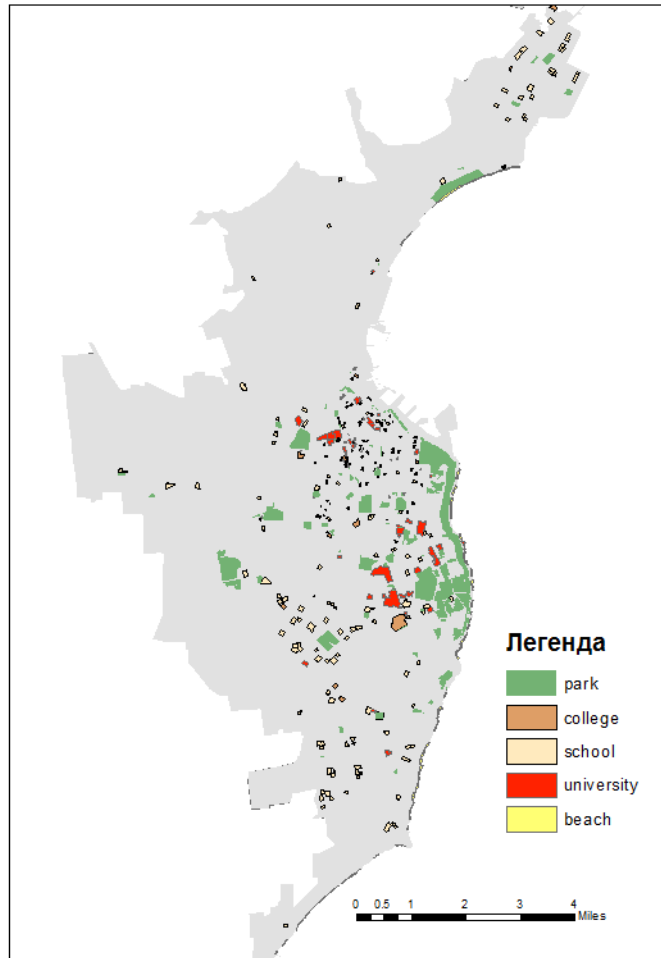


Рисунок 3.4 – Шар «Наявність покупців»

Для створення композитної карти евклідових відстаней до покупців виконаємо стандартизацію растрових карт, наведених на рис.3.5, у діапазоні $[0;1]$ з використанням апарату нечіткої логіки. Будемо вважати, що найкращою відстанню є відстань менша за 500 м, найгіршою – більше 1000 м. Таким чином, перекласифікуємо значення комірок растру відповідно монотонно спадної лінійної функції належності так, що коміркам зі значенням не більше 500 привласнимо значення 1, а коміркам зі значенням 1000 м і більше – 0. Всі проміжні значення монотонно спадають від 1 до 0. Стандартизовані

растрові карти евклідових відстаней для об'єктів з наявністю потенційних покупців наведені на рис.3.5.

Стандартизація критеріїв в діапазоні $[0,1]$ була виконана за допомогою інструменту Reclassify бібліотеки Spatial Analyst пакету ArcGIS 10.2, який виконує перекласифікацію растра в інтервалі $[0,10]$, а далі з використанням інструменту «Алгебра карт» усі значення комірок отриманого растру діляться на 10.

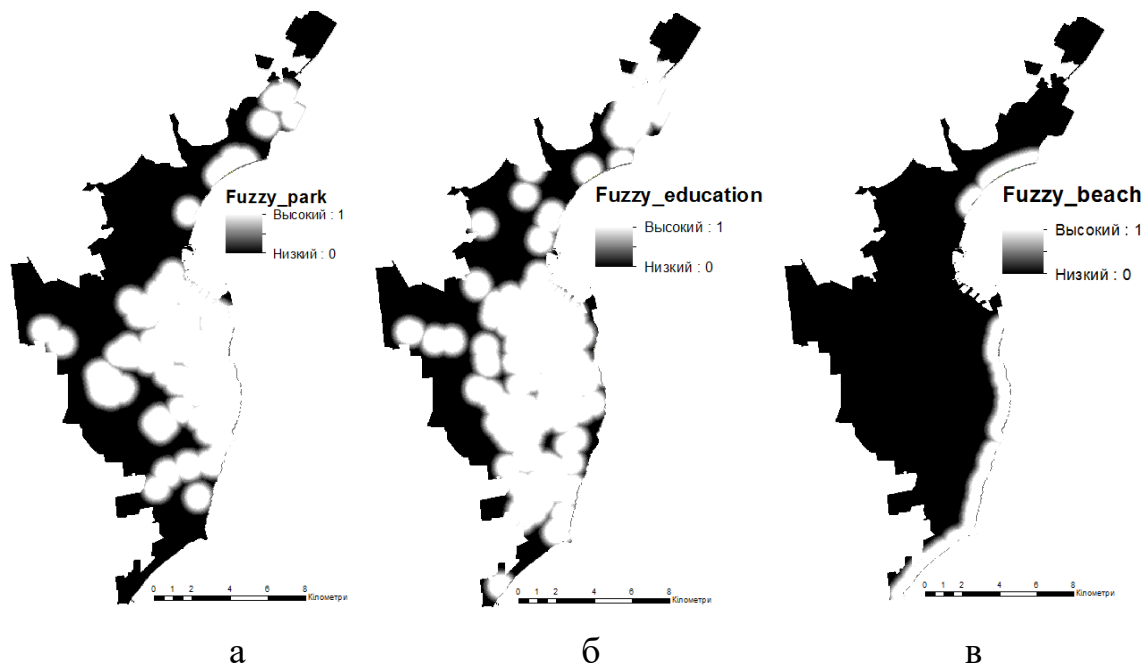


Рисунок 3.5 – Карти нечіткої стандартизації критеріїв:

а – паркові зони; б – навчальні заклади; в – пляжні зони

Для аналізу транспортної доступності за допомогою SQL запитів до бази даних (рис.3.6) були знайдені і об'єднані в окремі шари об'єкти головних автошляхів і пішохідних доріг, а також зупинки автотранспорту.

Отримані шари наведені на рис.3.6. Евклідові відстані для кожного типу об'єкту транспортної доступності представлені на рис.Б.2.

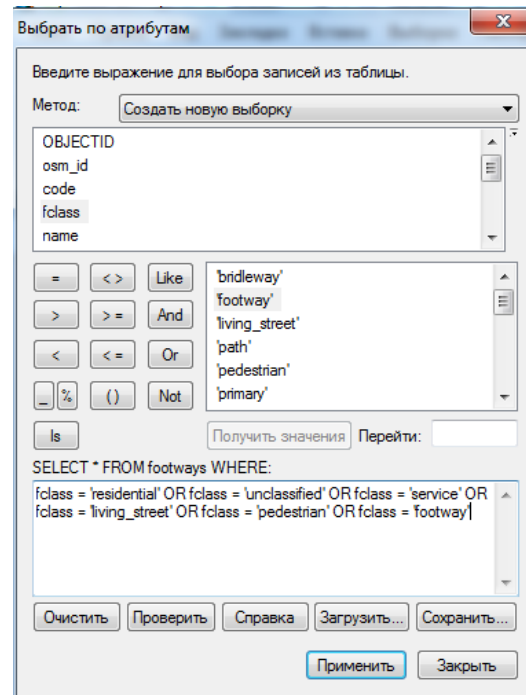


Рисунок 3.6 – SQL запит до бази даних для створення шару пішохідних доріг

Для стандартизації критеріїв транспортної доступності також будемо використовувати апарат нечіткої логіки. Оберемо лінійні функції належності (рис.3.7) і наступні контрольні точки (табл. 3.1):

- для пішохідних доріг найбільш підходить лінійна монотонно спадна функція належності з контрольними точками: $a=20$ м і $b= 50$ м. Чим ближче кав'ярня знаходиться до пішохідних доріг тим більша ймовірність, що її побачать покупці;

- для автобусних зупинок і головних магістралей оберемо симетричну лінійну функцію належності до нечіткої множини з контрольними точками: $a=20$ м, $b= 50$ м, $c = 100$ м, $d = 200$ м. Тобто найкраще місце розташування близько 50-100 м від зупинок і головних доріг.

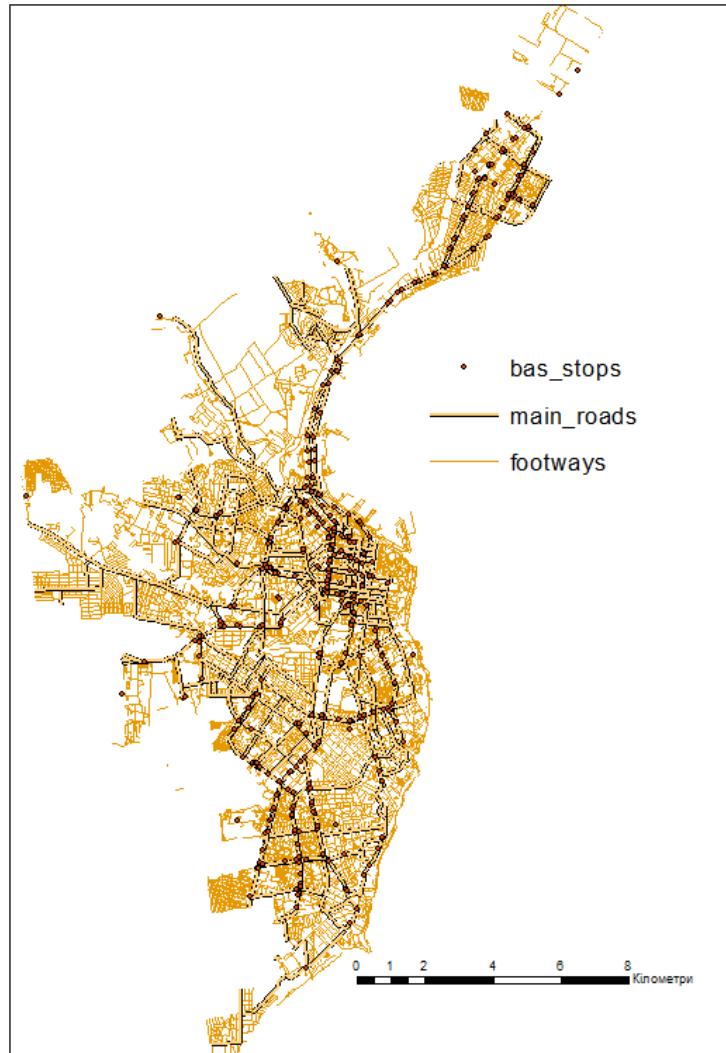


Рисунок 3.6 – Шар транспортної доступності

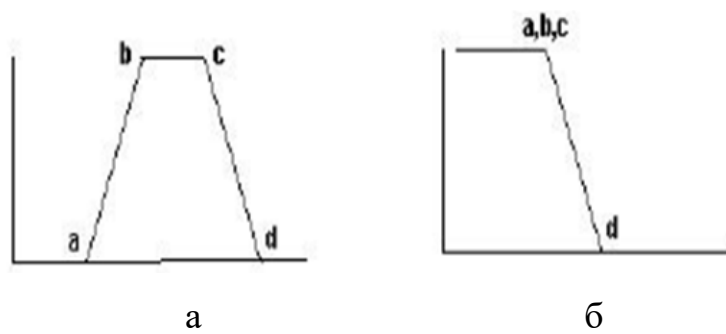


Рисунок 3.7 – Типи лінійних функцій належності:

а – симетрична; б – монотонно-спадна

Таблиця 3.1 – Нечіткі функції та множини з контрольними точками, що використовуються для критеріїв

Критерії	Контрольні точки				Тип функції належності	Форма нечіткої множини
	a	b	c	d		
Паркові зони (м)	500	500	500	1000	Лінійна	Монотонно спадна
Пляжні зони (м)	500	500	500	1000	Лінійна	Монотонно спадна
Навчальні заклади (м)	500	500	500	1000	Лінійна	Монотонно спадна
Автобусні зупинки (м)	20	50	100	200	Лінійна	Симетрична
Пішохідні дороги (м)	20	20	20	50	Лінійна	Монотонно спадна
Головні дороги (м))	20	50	100	200	Лінійна	Симетрична

Стандартизовані растрові карти евклідових відстаней для об'єктів з транспортної доступності (автобусних зупинок, пішохідних доріг та головних магістралей) наведені на рис.3.8. Стандартизація критеріїв в діапазоні [0,1] була виконана за допомогою інструменту Reclassify бібліотеки Spatial Analyst пакету ArcGIS 10.2, який виконує перекласифікацію растра в інтервалі [0,10], та інструменту «Алгебра карт», який ділить всі значення комірок отриманого растру на 10.

3.6 Визначення ваги критеріїв методом аналізу ієрархій

Визначення ваги критеріїв будемо виконувати за допомогою методу аналізу ієрархій, розробленим Томасом Сааті, з використанням спеціально розробленої програмної надбудови ArcGIS 10.2.

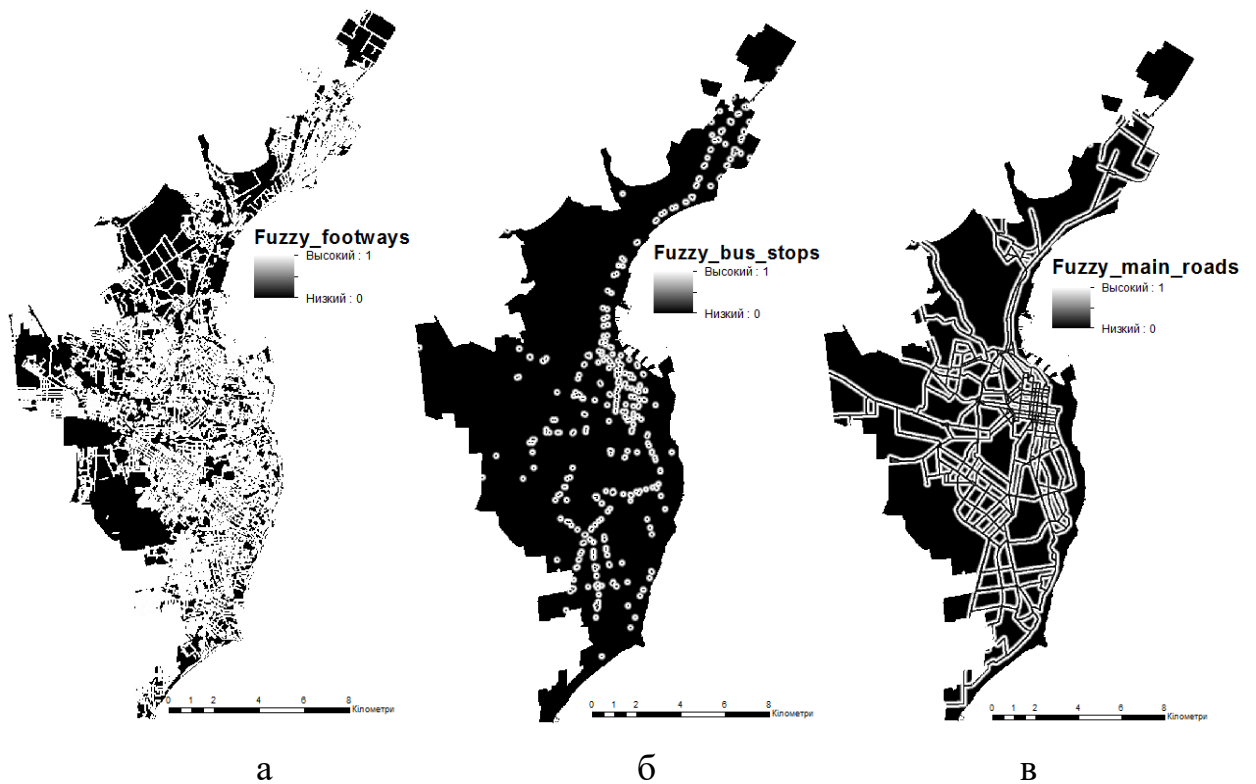


Рисунок 3.8 – Карты нечіткої стандартизації критеріїв:

а – пішохідні дороги; б – автобусні зупинки; в – головні магістралі

Надбудова являє собою окрему кнопку «АНР» на панелі інструментів пакету (рис.3.9) і дозволяє виконувати усі найважливі етапи розрахунку ваги критеріїв відповідно до алгоритму метода аналізу ієрархій.

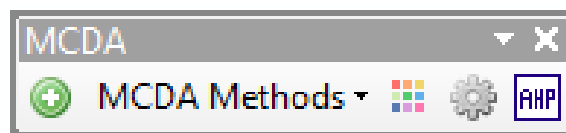


Рисунок 3.9 – Вигляд інтерфейсу надбудови «АНР»
на панелі інструментів

3.6.1 Побудова ієрархії критеріїв

На першому етапі на підставі існуючих у проекті карт будується ієрархія критеріїв, як це показано на рис. 3.10. Ієрархію складають 6 основних критеріїв: «Паркові зони», «Пляжні зони», «Навчальні заклади», «Автобусні зупинки», «Пішохідні дороги», «Головні магістралі».

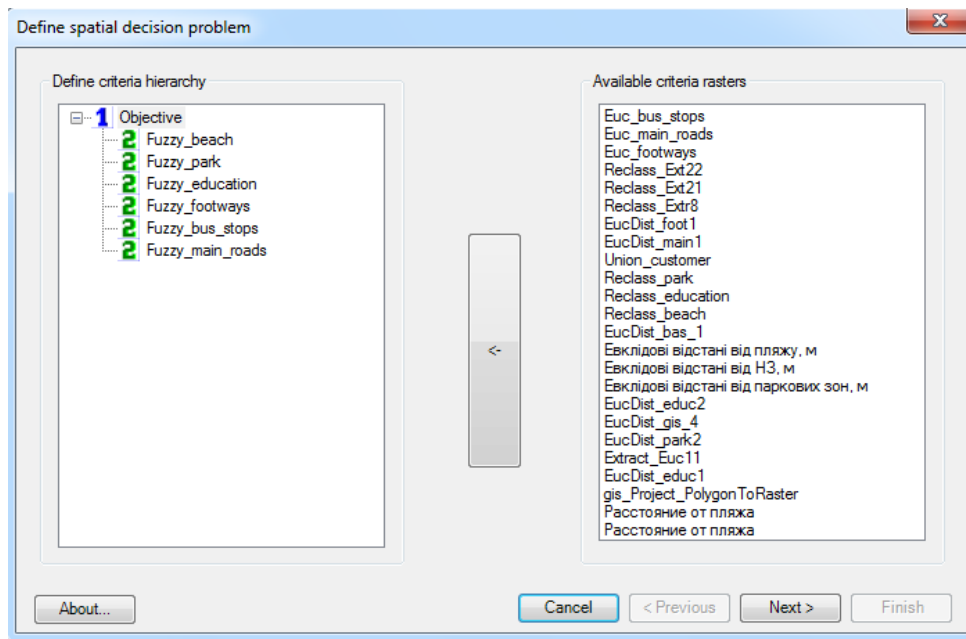


Рисунок 3.10 – Визначення ієрархії критеріїв

3.6.2 Побудова матриці попарних порівнянь

На другому етапі експерт виконує попарне порівняння важливості критеріїв, таким чином формується матриця попарних порівнянь (в нашому випадку розміром 6х6). Вікно введення оцінок експертів в надбудові показано на рис.3.11.

Для порівняння використовується фундаментальна шкалу абсолютних чисел Сааті [18], що приймає цілі значення від 1 до 9 (табл.3.2). Основною

перевагою цього способу порівнянь є те, що він є безрозмірним і не виникає проблем при приведенні до однакових одиниць виміру.

Правомірність цієї шкали доведена теоретично і практично при порівнянні з багатьма іншими відомими даними. Досвід показав, що при проведенні парних порівнянь, в основному, ставляться запитання: "Який з елементів є важливішим? Який найвірогідніший? Який з них найпривабливіший?". Ефективність такого способу проявляється лише при обмежованій кількості критеріїв (не більше 15).

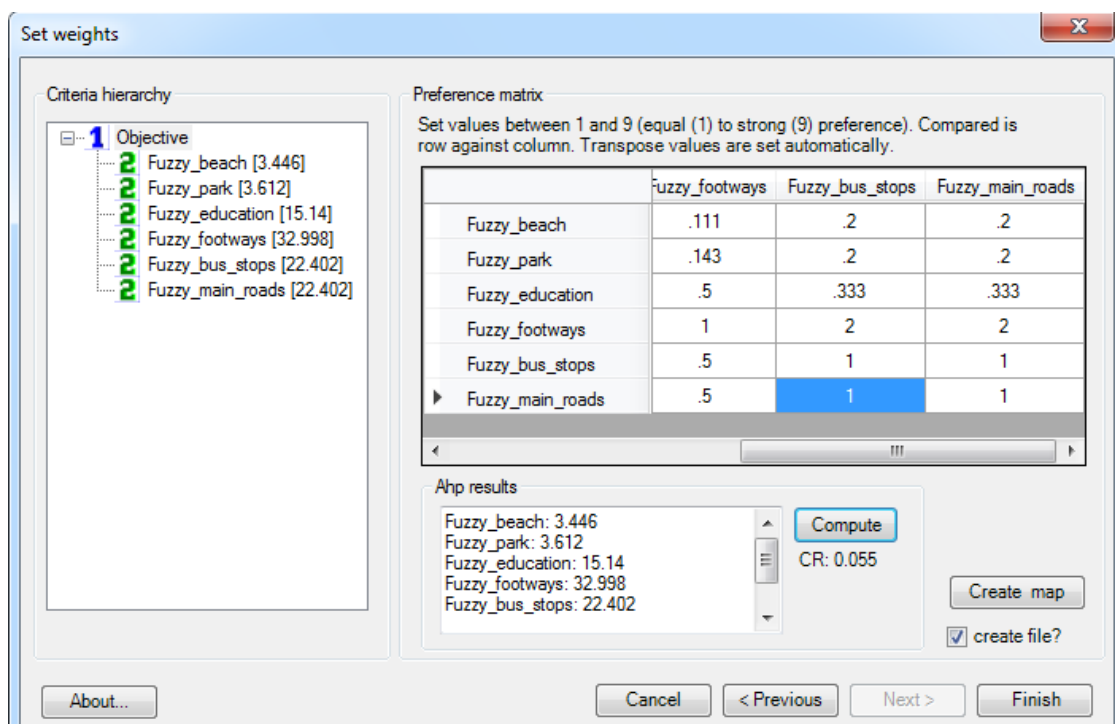


Рисунок 3.11 – Вікно розрахунку ваги критеріїв

Оцінки експертів базуються на їх досвіді і є суб'єктивними, для пом'якшення суб'єктивного характеру визначення важливості критеріїв доцільним є урахування думок декількох експертів з цієї проблеми. Таким чином окрема матриця будується для оцінок кожного експерта, а розраховані таким чином ваги далі усереднюються відповідно до кількості експертів.

Таблиця 3.2 – Фундаментальна шкалу абсолютних чисел Сааті

Відносна важливість	Визначення	Пояснення
1	однакова важливість	обидва елементи вносять однаковий вклад
3	один елемент трохи важливіший за другий	досвід дозволяє поставити один елемент трохи вище за другий
5	суттєва перевага	досвід дозволяє встановити безумовну перевагу одного над другим
7	значна перевага	один елемент настільки важливіший за другий, що є практично значимим
9	абсолютна перевага одного над другим	очевидність переваги підтверджується більшістю експертів
2,4,6,8	проміжні оцінки між сусідніми твердженнями	компромісне рішення
обернені величини чисел	якщо при порівнянні одного елемента з іншим, виставлено одне з чисел (1-9), то при зворотньому порівнянні матимемо обернену величину	

В деяких публікаціях вчені рекомендують застосовувати для визначення ваги критеріїв апарат нечіткої логіки, а саме трикутні нечіткі числа. Тоді попарна оцінка порівняння критеріїв виконується шляхом застосування нечіткої шкали з трикутними нечіткими числами (табл. 3.3). В даній роботі був використаний класичний алгоритм МАІ.

Матриця порівнянь, що була отримана в результаті оцінювання експерта приведена у додатку В.

Таблиця 3.3 – Фундаментальна шкалу Сааті для попарного порівняння з використанням нечітких чисел

Нечітка шкала	Визначення
(1,1,1)	однакова важливість
(2,3,4)	один елемент трохи важливіший за другий
(4,5,6)	суттєва перевага
(6,7,8)	значна перевага
(8,9,9)	абсолютна перевага одного над другим
($x-1, x, x+1$)	проміжні оцінки між сусідніми твердженнями

В результаті розрахунку власних векторів і власних значень матриці попарних порівнянь були отримані ваги кожного з критерію. Результати розрахунку наведені в табл. 3.4. Сумарне значення всіх ваг критеріїв дорівнює 100. Як видно, найбільш важливим критерієм експерт вважає наявність пішохідних доріг (32.998%), найменш важливим – пляжні зони (3.612 %).

Таблиця 3.4 – Результат розрахунку ваги критеріїв

№ з/п	Назва критерію	Вага, %
1	Пляжні зони	3,446
2	Паркові зони	3,612
3	Навчальні заклади	15,14
4	Пішохідні дороги	32,998
5	Автобусні зупинки	22,402
6	Головні магістралі	22,402

3.6.3 Перевірка узгодженості оцінки експертів

Для перевірки узгодженості оцінки експертів Томас Саати запропонував коефіцієнт узгодженості (CR). Розрахунок ступеня узгодженості виконується в два кроки. На першому кроці обчислюється індекс узгодженості:

$$CI = (\lambda_{max} - n)/(n - 1), \quad (3.1)$$

де n – матричний ранг і λ_{max} – максимальне власне значення матриці порівняння.

На другому кроці розраховується CR як відношення між CI до випадкового індексу (RI).

$$CR = \frac{CI}{RI} \quad (3.2)$$

Випадковий індекс (RI) залежить від ранжування матриці, а його значення розраховуються випадковою генерацією 500 матриць [18] (табл. 3.5).

Якщо коефіцієнт узгодженості (CR) менше або дорівнює 0.10, це доводить, що оцінки експерта були узгоджені і що немає необхідності змінювати оцінки [16]. Якщо CR перевищує 0.10, особі, що приймає рішення, слід повторити (або змінити) його оцінку, щоб поліпшити власну узгодженість.

Розрахунок показав, що для моделі коефіцієнт узгодженості $CR = 0.055$, що говорить про прийнятну узгодженість оцінок експерта, тобто повторювати процедуру порівняння не треба.

Таблиця 3.5 – Значення випадкового індексу (RI) в залежності від рангу матриці

n	1	2	3	4	5	6	7	8	9	10
-----	---	---	---	---	---	---	---	---	---	----

<i>RI</i>	0.00	0.00	0.52	0.89	1.11	1.25	1.35	1.40	1.45	1.49
-----------	------	------	------	------	------	------	------	------	------	------

3.7 Виконання зваженого накладання

В пакеті ArcGIS 10.2 є спеціальний інструмент «Зважене накладання» бібліотеки Spatial Analyst, який дозволяє виконати цю процедуру. Але його реалізація передбачає використання лише цілочисельних растрів і цілих значень ваги критеріїв. В нашому випадку, після нечіткої стандартизації значення атрибутів знаходяться в діапазоні [0,1], а ваги критеріїв є дійсними числами, тому будемо використовувати інший спосіб агрегації, а саме використовувати калькулятор растрів. Процес виконання зваженого накладання методом WLC в калькуляторі растра показаний на рис.3.12.

Остаточна карта придатності, що була отримана в результаті виконання операції зваженого накладання і після перекласифікації, показана на рис.3.13.

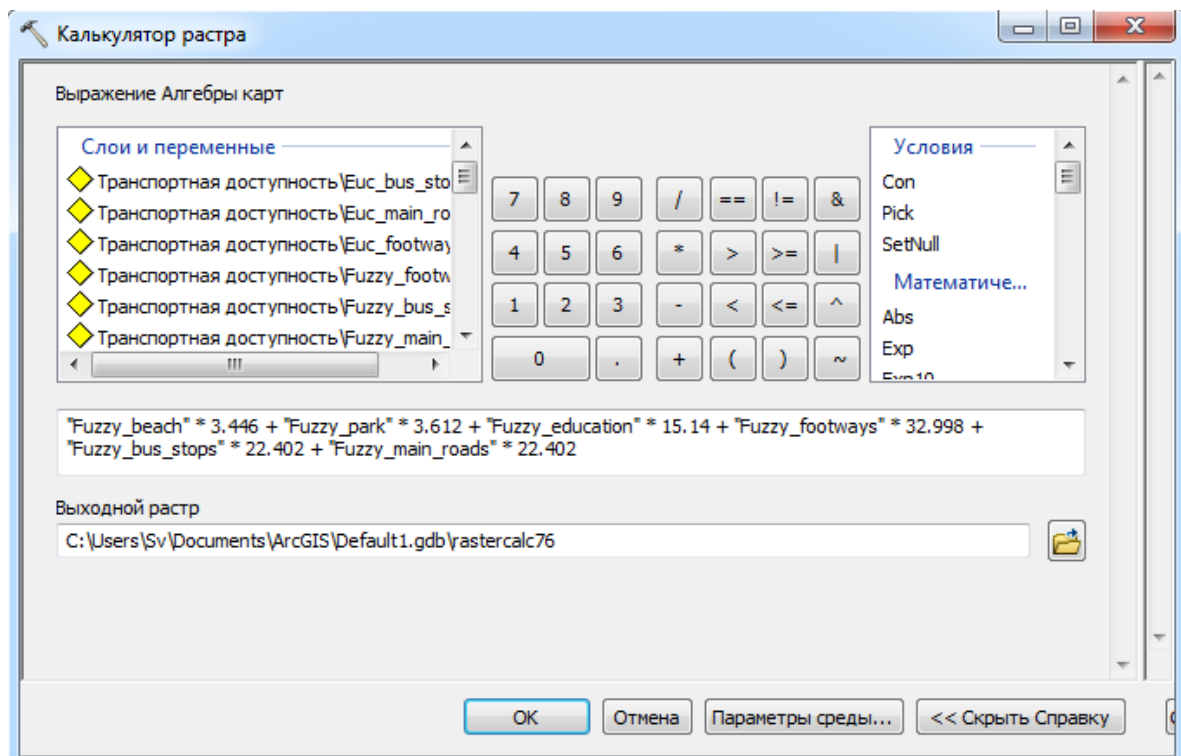


Рисунок 3.12 – Виконання зваженого накладання з використанням калькулятора растра

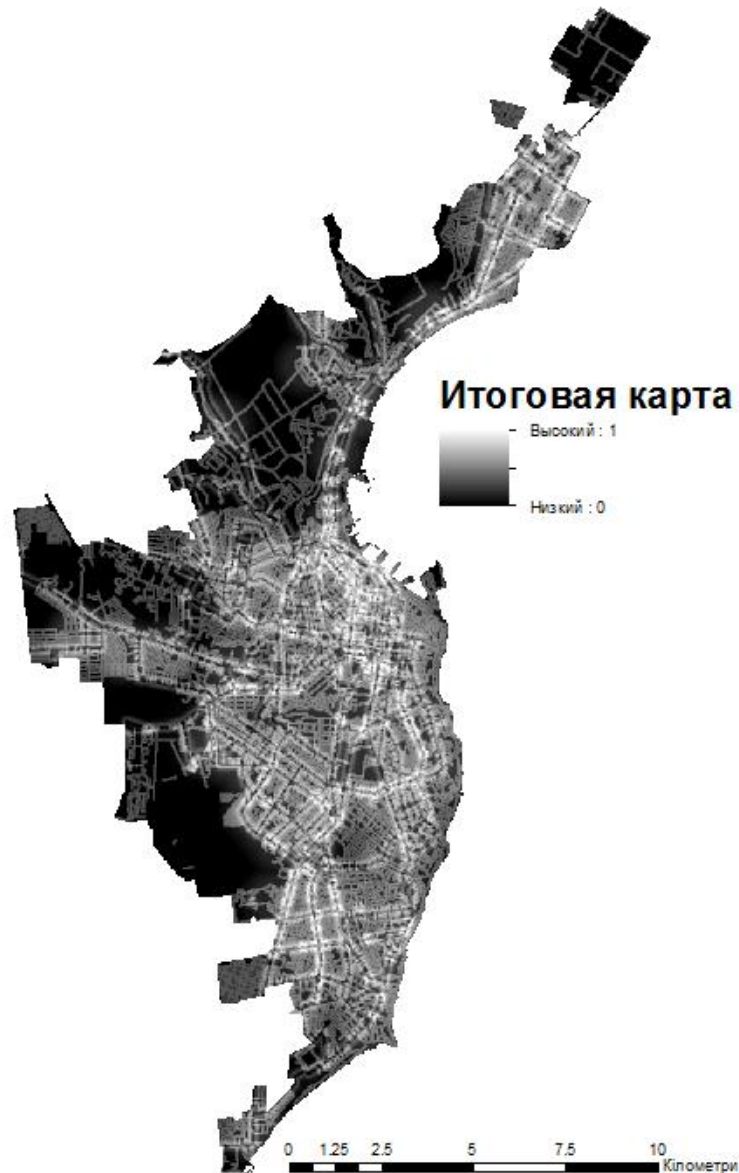


Рисунок 3.13 – Карта придатності, яка отримана в результаті виконання операції зваженого накладання

Як можна побачити на карті придатності (рис.3.13) найбільш придатні ділянки під кав'ярню розташовані вздовж морського побережжя. Найбільш таких місць в центрі міста. Це пов'язано з наявністю там більшої кількості навчальних закладів і зон відпочинку.

Далі з цієї карти за допомогою кон'юнктивного фільтру можна вилучити зони на яких немає будівель. Таким чином будуть враховані обмеження. На карті, що зображена на рис.3.14 показаний фрагмент підсумкової карти

придатності з шаром будівель, як бачимо зображення цих двох шарів разом дозволяє ЛПР краще побачити будівлі, які мають високу цінність і прийняти відповідне рішення.

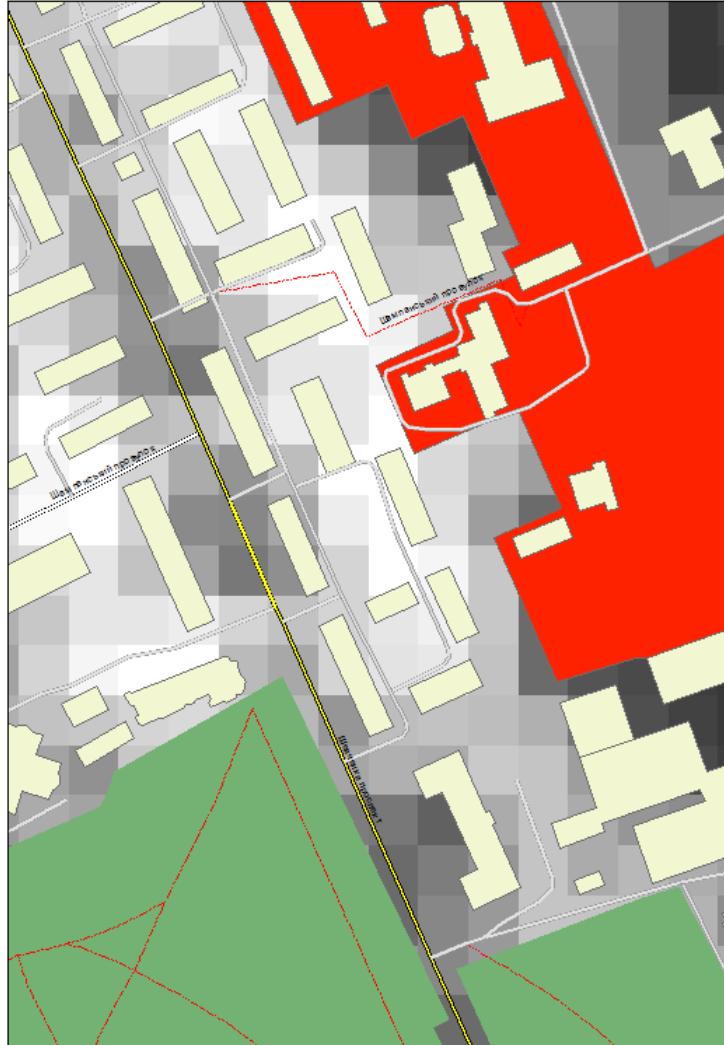


Рисунок 3.14 – Зображення шару будівель на підсумковій карті придатності

3.8 Розробка моделі геообробки даних в середовищі ModelBuilder

Конструктор моделей ArcGIS (Model Builder) – потужний засіб, що дозволяє, за допомогою побудови відповідних моделей більш ефективно управляти процесами обробки даних в ArcGIS. Елементом моделі може бути один з інструментів з ArcToolbox, скрипт на мові Python, VBScript та ін.

Модель в ModelBuilder – це відображення опису робочих процесів, які з'єднані один з одним, в послідовності інструментів геообробки, подаючи вихід одного інструменту на вхід іншого інструменту. ModelBuilder можна також розглядати як візуальну мову програмування для побудови робочих потоків.

ModelBuilder має три набори параметрів, що використовуються в моделі:

1) Властивості моделі – ці властивості, які дозволяють задавати назву моделі, напис, опис, відносний шлях, властивості параметрів, змінні середовища моделі, довідку і кількість проходів.

2) Властивості діаграми – це властивості, що дозволяють змінювати розташування елементів на діаграмі, а також її колір і стиль.

3) Властивості відображення – це властивості, що дозволяють змінювати зовнішній вигляд і інші графічні властивості окремих елементів.

Елементи моделі є основними блоками, з яких будуються моделі. Існує три типи елементів моделі:

1) Інструменти – є основними блоками, з яких будуються робочі процеси в моделі. З їх допомогою виконуються різні операції з географічними або табличними даними. Коли інструменти додаються в модель, вони стають елементами моделі.

2) Змінні – елементами моделі, в яких зберігаються значення або посилання на дані. Існує два типи змінних:

– змінні даних – є елементами моделі, в яких містяться описові відомості про дані, що зберігаються на диску. В них описуються такі властивості даних, як відомості про поле, просторова прив'язка і шлях;

– змінні значень – це такі значення як рядки, числа, логічні значення (true/false), просторові прив'язки, лінійні одиниці або екстенти. Змінні значень можуть містити що завгодно, крім посилань на дані, що зберігаються на диску.

3) З'єднувачі – з'єднують дані і значення з інструментами. Стрілки з'єднувачів вказують напрямок обробки. Існує чотири типи з'єднувачів.

- дані – з'єднують змінні даних і значень з інструментами;
- середовище – з'єднують змінну, що містить параметр середовища (дані або значення), з інструментом;
- попередня умова – з'єднує змінну з інструментом. Інструмент буде виконуватися тільки після створення вмісту змінної попередньої умови;
- зворотній зв'язок – направляє вихідні дані інструменту назад в той же інструмент в якості вхідних даних.

На рис.3.15 показано класифікація елементів моделі в ModelBuilder.

При запуску процесів з ModelBuilder їх виконання можна відстежувати в діалоговому вікні ходу геообробки. При запуску моделі відбувається наступне: перевіряються всі змінні моделі і виконуються інструменти, які мають стан «Готовий до виконання» (Ready-to-run). Якщо будь-які вихідні змінні повинні бути додані до зображення, вихідні дані додаються в таблицю змісту ArcMap. Інструменти, які не мають стану «Готовий до виконання» (Ready-to-run) або залежать від інструменту, не готового до виконання, не виконуються.

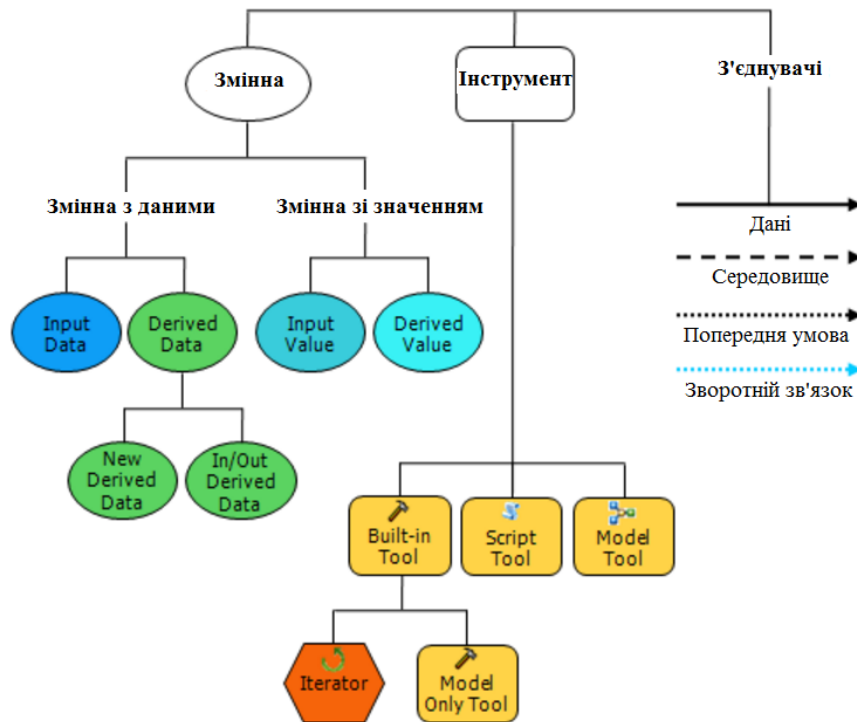


Рисунок 3.15 – Класифікація елементів моделі в ModelBuilder

ModelBuilder дозволяє більш ефективно створювати прості або складні моделі за допомогою ряду методів. Завдяки цьому можна більш ефективно управляти логічною послідовністю виконання моделі, її даними, а також швидше здійснювати завдання, що повторюються. Підтримуються:

- ітерації або циклічна організація, що забезпечує кількаразове повторення процесу;
- списки з декількома значеннями;
- зворотний зв'язок, коли в якості вхідних даних використовуються вихідні дані;
- попередні умови для управління послідовністю операцій в моделі;
- логіка якщо-інакше (if-then-else), що дозволяє виконувати різні дії в різних умовах.

Інструменти моделі повністю інтегровані в середовище геообробки і працювати з ними можна так само, як з системними інструментами або інструментами скриптів. Як і всі інші інструменти геообробки, інструменти

моделей можна запускати з діалогового вікна, через скрипти Python, а також додавати і запускати їх в іншій моделі. Скрипти Python і зовнішні програми, що запускаються за допомогою скрипта, можна інтегрувати в модель так само, як і системні інструменти, що додаються і запускаються в моделі.

Процес розробки моделі геообробки даних в середовищі ModelBuilder проілюструємо на прикладі створення нечіткого шару для підкласу «Наявність покупців». В процесі обробки були задіяні наступні інструменти: вибір за атрибутом, проєціювання, евклідова відстань, перекласифікація та калькулятор растру. Значення ваг для критеріїв додаються як параметри. Вигляд моделі наведений на рис.3.16.

На рис.3.17 представлено вікно запуску цієї моделі, у якому треба вказати ваги для критеріїв пляжні зоні, паркові зоні і навчальні заклади, а також вказати базу геоданих в яку буде збережений результат роботи моделі.

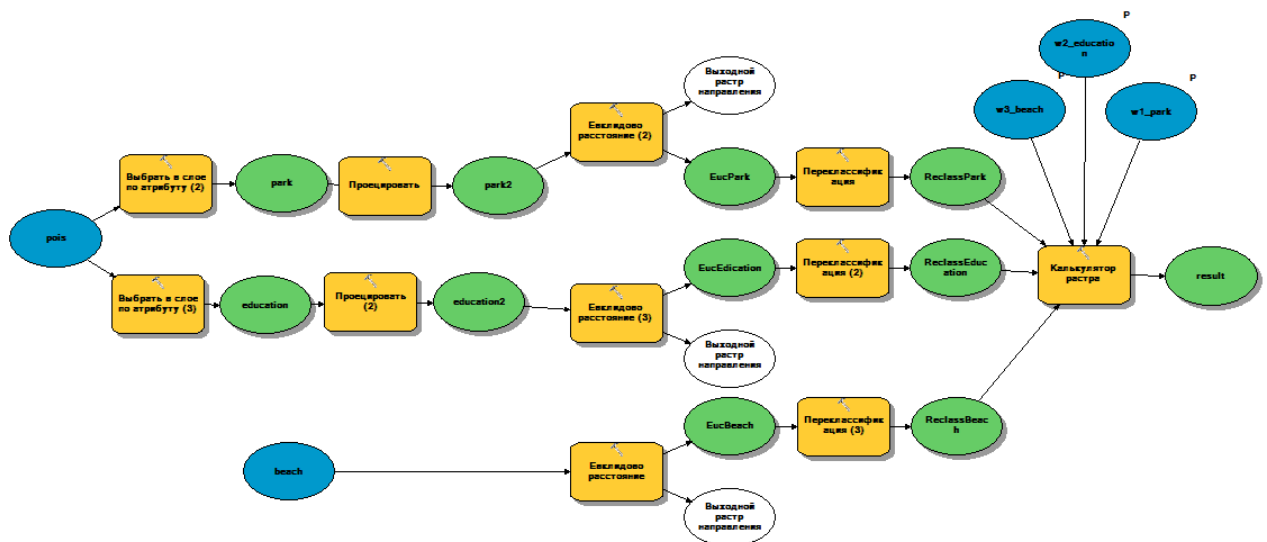


Рисунок 3.16 – Модель побудови нечіткого шару для підкласу «Наявність покупців»

Повна модель геообробки даних системи підтримки прийняття рішення про розміщення кав'ярні в м. Одесі наведена у додатку Г.

В моделі задіяні 4 шари карти OpenStreetMap з яких за допомогою інструменту «Вибрати за атрибутом» сформовано 6 карт критеріїв: пляжні зони, паркові зони, навчальні заклади, автобусні зупинки, пішохідні дороги, головні магістралі. Всі карти критеріїв було перепроєцровано до прямокутної системи координат, а далі за допомогою інструменту «Евклідова відстань» було створено растрові шари критеріїв, які містять значення відстаней від заданих цілій в метрах.

Отримані растрові шари було перекласифіковано відповідно до заданих функцій приналежностей до нечіткої множини і за допомогою методу WLC і інструменту «Калькулятор растра» агреговане в підсумковий растр придатності.

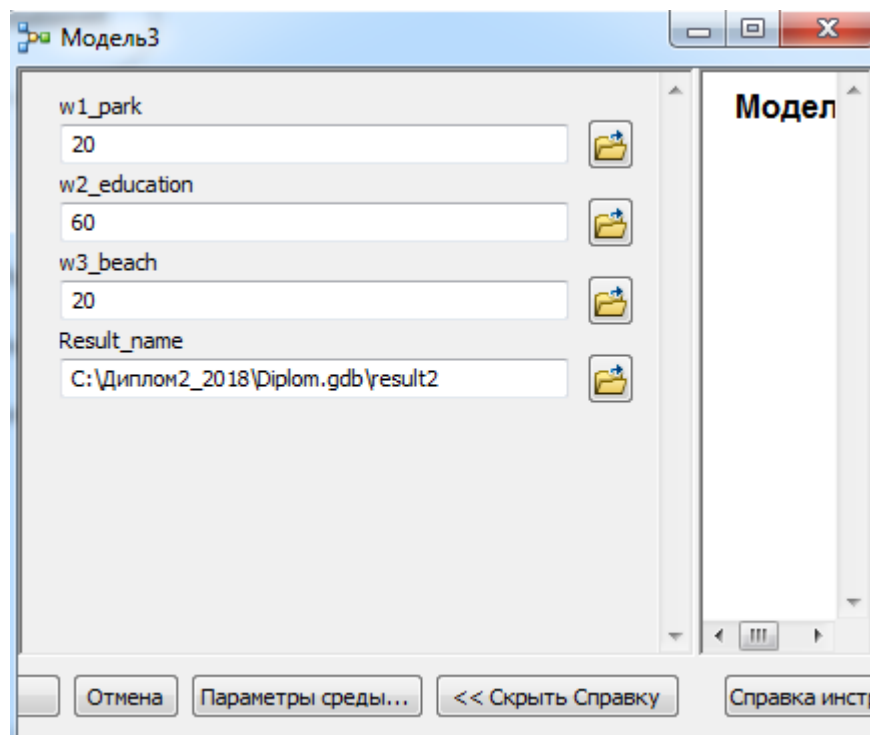


Рисунок 3.17 – Вікно запуску моделі

Для виконання підсумовування перекласифікованих растрових шарів критеріїв використовуються значення їх ваг, які були підраховані за допомогою окремого скрипта АНР. Ваги додаються до моделі у вигляді вхідних параметрів.

4 СТВОРЕННЯ ІНТЕРАКТИВНОЇ КАРТИ

В останні роки розвиток Internet-технологій все більшою мірою стало дозволяти користувачам мережі використовувати нові засоби доставки інформації, в тому числі і просторової.

Web-картографія – це область комп'ютерних технологій, пов'язана з доставкою просторових даних кінцевому користувачеві. безумовно, web-картографія є одним з напрямків геоінформаційних технологій в цілому. Основними завданнями web-картографії є:

- 1) Візуалізація існуючої інформації – просторове передання інформації.
- 2) Полегшення роботи з просторовою інформацією в мережі Internet, пошук, прокладка маршрутів і інші послуги, засновані на місцезнаходження об'єктів (LBS – location based services).

Розглянемо технічні засоби які можуть бути використані для створення на базі існуючих в проекті карт інтерактивної карти для сумісної роботи декількох користувачів. І обґрунтуємо вибір технічних засобів для створення власної інтерактивної карти.

4.1 Веб-сервер GeoServer

GeoServer – це сервер з відкритим вихідним кодом, що дозволяє легко управляти ГІС-даними [26]. Він надає дані у відповідності зі стандартами WMS і WFS. Він також підтримує сервіс WFS-T для модифікації актуальних даних, а також мову географічної розмітки (GML), мову розмітки Keyhole (KML) і багато інших форматів. Дані організовуються в робочі області та сховища даних, при цьому використовується дискова система зберігання або PostGIS. За допомогою вбудованого REST API можна управляти наборами даних в режимі реального часу. Це ПЗ також включає в себе готовий Web-інтерфейс.

WMS – це інтерфейс для запитуваних зображень карти, а WFS – інтерфейс для запитуваних векторних даних карти. WMS частіше використовується в ситуаціях, коли метою є виключно відображення даних. WFS зручніше при взаємодії користувачів з вихідними даними (наприклад, запит значень таблиці атрибутів або зміна даних). Модифікацію WFS часто називають WFS-T або WFS з транзакціями.

В якості системи зберігання даних з GeoServer найчастіше використовують СКБД PostgreSQL з розширенням PostGIS. А для створення карт на стороні клієнта потрібно OpenLayers – відмінний JavaScript-інструмент, який дозволяє розробляти (в тісній взаємодії з GeoServer) Web-додатки для представлення просторових даних.

Потреби в ресурсах безпосередньо залежать від обсягу трафіку, який планується направляти на GeoServer. На диску повинно бути достатньо місця для зберігання ГІС-даних і близько 5 ГБ для операційної системи і GeoServer.

Рекомендується принаймні 512 МБ оперативної пам'яті. Можливе використання операційних систем Microsoft Windows і Mac OS, або Linux.

Для GeoServer існує безліч розширень, які дозволяють використовувати більше джерел даних (наприклад, IBM DB2) і нові функції (наприклад, Web-сервіс геопроцесінга (WPS) і виведення в Excel-форматі). Ці розширення можна знайти на сторінці завантаження GeoServer.

Після установки GeoServer стає доступним його адміністративний інтерфейс для виконання налаштування (рис.4.1).

GeoServer організовує дані (шари) в робочі області та сховища даних. Шар – це індивідуальний набір даних, наприклад дороги або озера. Робоча область забезпечує логічне групування даних для конкретного проекту або географічного регіону. Робоча область може містити одне або кілька сховищ даних, що визначають фізичне місце розташування даних (наприклад, каталог або файл на диску або таблиця бази даних PostGIS).

The screenshot shows the GeoServer administrative interface. The main content area is titled 'Слои' (Layers) and contains a table of layers. The table has the following columns: Тип (Type), Заголовок (Title), Название (Name), Хранилище (Storage), Включено (Enabled), and Родной SRS (Native SRS). The table lists 20 layers, including administrative boundaries, world rectangle, Manhattan points of interest, roads, and various sample data sets. The interface also includes a sidebar with navigation options and a top navigation bar with the GeoServer logo and a 'Вышли' (Logout) button.

Тип	Заголовок	Название	Хранилище	Включено	Родной SRS
<input type="checkbox"/>	Административные границы	LessonG:boundary-polygon	Ukraine	✓	EPSG:4326
<input type="checkbox"/>	World rectangle	tiger:giant_polygon	nyc	✓	EPSG:4326
<input type="checkbox"/>	Manhattan (NY) points of interest	tiger:poi	nyc	✓	EPSG:4326
<input type="checkbox"/>	Manhattan (NY) landmarks	tiger:poly_landmarks	nyc	✓	EPSG:4326
<input type="checkbox"/>	Manhattan (NY) roads	tiger:tiger_roads	nyc	✓	EPSG:4326
<input type="checkbox"/>	A sample ArcGrid file	nurc:Arc_Sample	arcGridSample	✓	EPSG:4326
<input type="checkbox"/>	North America sample imagery	nurc:Img_Sample	worldImageSample	✓	EPSG:4326
<input type="checkbox"/>	PK50095	nurc:PK50095	img_sample2	✓	EPSG:32633
<input type="checkbox"/>	mosaic	nurc:mosaic	mosaic	✓	EPSG:4326
<input type="checkbox"/>	USA Population	topp:states	states_shapefile	✓	EPSG:4326
<input type="checkbox"/>	Tasmania cities	topp:tasmania_cities	taz_shapes	✓	EPSG:4326
<input type="checkbox"/>	Tasmania roads	topp:tasmania_roads	taz_shapes	✓	EPSG:4326
<input type="checkbox"/>	Tasmania state boundaries	topp:tasmania_state_boundaries	taz_shapes	✓	EPSG:4326
<input type="checkbox"/>	Tasmania water bodies	topp:tasmania_water_bodies	taz_shapes	✓	EPSG:4326
<input type="checkbox"/>	Spearfish archeological sites	sf:archsites	sf	✓	EPSG:26713
<input type="checkbox"/>	Spearfish bug locations	sf:bugsites	sf	✓	EPSG:26713
<input type="checkbox"/>	Spearfish restricted areas	sf:restricted	sf	✓	EPSG:26713
<input type="checkbox"/>	Spearfish roads	sf:roads	sf	✓	EPSG:26713
<input type="checkbox"/>	Spearfish elevation	sf:sfdem	sfdem	✓	EPSG:26713
<input type="checkbox"/>	Spearfish streams	sf:streams	sf	✓	EPSG:26713

Рисунок 4.1 – Адміністративний інтерфейс GeoServer

Після підготовки даних можна обрати відповідні варіанти в меню Data, щоб додати їх в GeoServer. Для більшості опцій можна залишити значення за замовчуванням. Додавання шарів називається публікацією шарів. Щоб виконати публікацію, виберіть потрібну систему координат (SRS - система просторової прив'язки) і натисніть Compute під кожним з двох обмежують прямокутників. Повторіть цю процедуру для всіх вихідних робочих областей, сховищ даних і шарів.

При використанні PostGIS, для досягнення найкращих результатів слід імпортувати хоча б один шейп-об'єкт в кожену таблицю бази даних, перш ніж розмістити її на GeoServer. Після настройки шарів можна переглянути їх в GeoServer, як показано на рис.4.2.



Рисунок 4.2 – Перегляд шару даних в GeoServer

GeoServer має API, який дозволяє виконувати більшість операцій програмним способом. Цей API робить GeoServer корисним інструментом

управління даними і основою для комплексного Web-додатки, яке автоматично додає і видаляє шари по мірі необхідності. GeoServer API можна використовувати безпосередньо, але краще створити інтерфейсний PHP-клас, який спрощує операції і інкапсулює необхідні взаємодії.

Всі операції API виконуються за допомогою запитів GET, POST, PUT або DELETE по URL-адресою `http://вашсервер:8080/geoserver/rest/`. Таким чином, запит GET за адресою `http://вашсервер:8080/geoserver/rest/workspaces.xml` повертає всі робочі області в форматі XML. У загальному випадку запит GET повертає результати в тому ж форматі, який використовувався при створенні екземпляра об'єкта за допомогою запиту POST або при його оновленні за допомогою запиту PUT. Лише в рідкісних випадках доведеться вказувати всі деталі, які повертаються запитом GET.

4.2 База даних PostgreSQL/PostGIS

PostGIS – розширення об'єктно-реляційної СУБД PostgreSQL призначене для зберігання в базі географічних даних [27]. PostGIS включає підтримку просторових індексів R-Tree/GiST і функції обробки геоданих. Використання для збереження геоданих СУБД, а не окремих шейп-файлів дозволяє організувати сумісну роботу великої кількості користувачів, як це показано на рис.4.3.

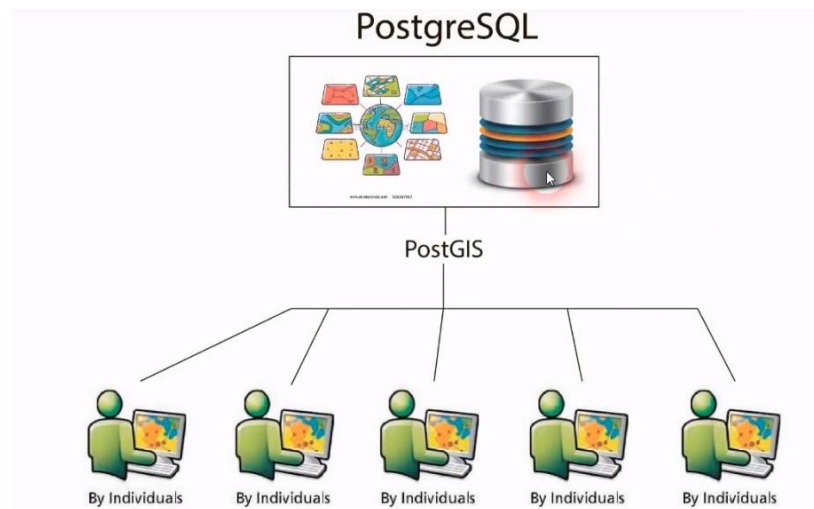


Рисунок 4.3 – Використання СКБД PostgreSQL для сумісного редагування геоданих багатьма користувачами

При створенні просторової бази даних автоматично створюються дві таблиці метаданих – `SPATIAL_REF_SYS` і `GEOMETRY_COLUMNS`. Вони створюються відповідно до специфікації «Open Geospatial Consortium Simple Features for SQL specification», випущеної OGC і описує стандартні типи об'єктів ГІС, функції для маніпуляції ними і набір таблиць метаданих.

Таблиця `GEOMETRY_COLUMNS` зберігає інформацію про таблиці бази даних, що містять просторову інформацію. Її заповнення здійснюється вручну, або як наслідок виконання спеціальної процедури OGC `AddGeometryColumn ()`.

Таблиця `SPATIAL_REF_SYS` містить числові ідентифікатори і текстові описи систем координат, що використовуються в просторової базі даних. Одним з полів цієї таблиці є поле `SRID` – унікальний ідентифікатор, однозначно визначає систему координат. `SRID` вдає із себе числовий код, якому відповідає певна система координат. Наприклад, поширений код EPSG 4326 відповідає географічній системі координат WGS84.

В роботі ми будемо наповнювати нашу просторову базу даних імпортувавши геодані з share-файлів, які будуть створені у QGIS. Для цього можна

спочатку перетворити дані з shape-файлу в формат зрозумілий PostgreSQL за допомогою утиліти shp2pgsql, а потім завантажити їх в базу даних. Можна завантажувати дані безпосередньо в пакеті QGIS, при цьому всі зміни будуть моментально відображені в адміністративному інтерфейсі СКБД PostgreSQL (рис.4.5).

Бібліотека OpenLayers дозволяє дуже швидко і легко створити web-інтерфейс для відображення картографічних матеріалів, представлених в різних форматах і розташованих на різних серверах. Завдяки OpenLayers розробник має можливість створити, наприклад, власну карту, що включає шари, що надаються WMS (і WFS) серверами, такими як Mapserver, ArcIMS або Geoserver, і даними картографічних сервісів Google. Бібліотека є розробкою з відкритим вихідним кодом і розробляється за спонсорської підтримки проекту MetaCarta, який використовує OpenLayers в своїх розробках. Проте, OpenLayers є незалежним вільно поширювальним продуктом.

Крім власне візуалізації WMS і WFS шарів на єдиної web-карті, а також редагування наданих даних (тільки для WFS-T серверів) OpenLayers володіє наступними можливостями:

- додавання на карту панелі навігації (включена за замовчуванням). На панелі знаходяться кнопки зсуву карти (північ-південь, захід-схід), збільшення і зменшення масштабу;
- зсуву карти за допомогою миші;
- зміни масштабу карти при прокручуванні середнього колеса миші;
- отримання координат точки, над якою знаходяться покажчик миші;
- додавання панелі управління видимістю / невидимістю шарів карти;
- вибору довільного об'єкта і отримання атрибутивної інформації про нього;
- управління прозорістю використовуваних шарів карти;
- додавання до карти визначаються користувачем елементів (точок, ліній, полігонів).

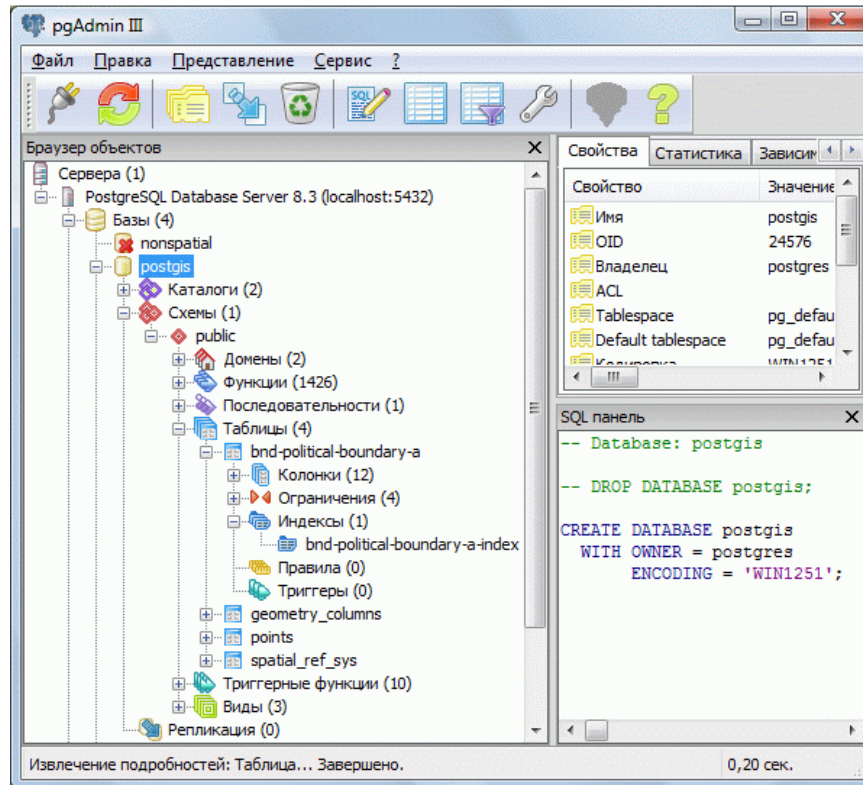


Рисунок 4.5 – Адміністративна панель СКБД PostgreSQL

4.3 Бібліотека OpenLayers

Безліч інших можливостей OpenLayers, які, разом з прикладами їх використання і документацій доступні на офіційному сайті [28].

У OpenLayers API є два основних поняття, ідентичним аналогічним пропозицій Mapserver: це об'єкти Map і Layer. Map зберігає інформацію про проекції, географічним охопленням, одиницях виміру та інших параметрах карти в цілому. У середині карти (Map) дані задаються за допомогою одного або декількох об'єктів Layer. Layer містить інформацію про шари даних, які будуть поміщені на карту, і про те, як кожен з цих шарів повинен відображатися на карті.

ВИСНОВКИ

У магістерській роботі був виконаний аналіз придатності за допомогою мультикритеріального аналізу рішень і інструментів просторового моделювання ГІС. Вирішена проблема пошуку оптимального місця розташування кав'ярні в м.Одеса.

Показано, що обґрунтованим є використання просторового моделювання і методів прийняття рішень в інтелектуальному аналізі даних. Інтеграція даних методів дозволяє вилучити знання з даних, особливо якщо ці дані мають просторову прив'язку.

Для виконання аналізу були використані дані картографічного сервісу OpenStreetMap. Шляхом рендеренгу отримані необхідні векторні шарі, що представляють собою об'єкти інфраструктури міста. За допомогою SQL запиту сформовані векторні шарі критеріїв: пляжні зони, паркові зони, навчальні заклади, пішохідні дороги, головні магістралі, автобусні зупинки. Для даних векторних шарів отримані растрові шарі евклідових відстаней, які далі були стартизовані за допомогою апарату нечіткої логіки відповідно до заданих функцій приналежностей.

Зважування критеріїв було виконано методом аналізу ієрархій. Для цього створена матриця попарних порівнянь, на основі оцінювання критеріїв експертом за фундаментальною шкалою абсолютних чисел Сааті. Далі розраховані власні вектора та власні числа цієї матриці на основі яких були отримані ваги критеріїв. Сумарне значення всіх ваг критеріїв дорівнює 100. Коефіцієнт узгодженості $CR = 0.083$, що говорить про прийнятну узгодженість оцінок експерта. Для виконання етапу, пов'язаного з розрахунком ваги критеріїв згідно алгоритму МАІ використана окрема надбудова для пакету ArcGIS 10.2.

Агрегація карт була виконана методом зваженого накладання. В роботі була створена загальна модель геообробки даних системи підтримки прийняття рішення, яка враховує послідовність застосування всіх інструментів па-

кету від обробки та растерізації вихідних шарів критеріїв до їх перекласифікації і накладання. Модель виконана у візуальному середовищі ModelBuilder.

ПЕРЕЛІК ДЖЕРЕЛ ПОСИЛАННЯ

1. Worboys, M. and Duckham, M. GIS: A Computing Perspective. CRC press. 2004. P. 34–56
2. Goodchild, M. F. The current status of gis and spatial analysis. Journal of Geographical Systems. 2000, 2(1): P.5–10
3. Longley, P., Goodchild, M., Maguire, D., and Rhind, D. Geographic Information Systems and Science. John Wiley & Sons, 3 edition. 2010, 5(2): P.15–47
4. Heywood, I., Cornelius, S., and Carver, S. (2002). An Introduction to Geographical Information Systems, 2nd ed. Pearson Education Limited, Toronto. 2002, P.156–164
5. Офіційний сайт компанії ESRI Inc. [Електронний ресурс]. – Режим доступу: www.esri.com
6. Сайт справочної системи ArcGIS. [Електронний ресурс]. – Режим доступу: webhelp.esri.com/arcgisdesktop
7. ArcGIS Resources [Електронний ресурс] – Режим доступу: <http://resources.arcgis.com/en/home/>
8. Open Source Initiative [Електронний ресурс] – Режим доступу: <https://opensource.org/>
9. FreeGIS.org. [Електронний ресурс] – Режим доступу: <http://freegis.org/>
10. Open Source Geospatial Foundation. [Електронний ресурс] – Режим доступу: <http://osgeo.org>
11. Open Geospatial Consortium. [Електронний ресурс] – Режим доступу: <http://opengeospatial.org>
12. Мелихов А.Н., Берштейн Л.С., Коровин С.Я. Ситуационные советующие системы с нечеткой логикой. – М.: Наука, 1990. – 272 с.
13. Трахтенгерц Э.А. Компьютерная поддержка принятия решений: Научно-практическое издание. – М.: СИНТЕГ, 1998. – 376 с.

14. Берштейн Л.С., Карелин В.П., Целых А.Н. Модели и методы принятия решений в интегрированных интеллектуальных системах. – Ростов-на-Дону: Изд-во РГУ, 1999. – 276 с.
15. Belton, V. and Stewart, T. Multiple Criteria Decision Analysis: An Integrated Approach. Kluwer Academic Publishers. 2002, P.78–97
16. Janssen, R. On the use of multi-criteria analysis in environmental impact assessment in the netherlands. *Journal of Multi-Criteria Decision Analysis*, 2001, 10(2): P.101–109.
17. Marler, R. T. and Arora, J. S. The weighted sum method for multi-objective optimization: new insights. *Structural and multidisciplinary optimization*. 2010, 41(6): P. 853–862.
18. Saaty, T. *The Analytic Hierarchy Process*. McGraw-Hill, New York. 1980. P.47
19. Saaty, T. Decision making with the analytic hierarchy process. *Scientia Iranica*. 2002, 9(3): P. 215–229.
20. Saaty, T. The analytic hierarchy and analytic network processes for the measurement of intangible criteria and for decision-making. In Figueira, F., Greco, S., and Ehrgott, M., editors, *Multiple Criteria Decision Analysis: State of the Art Surveys*, 2005, P. 345–408.
21. Заде Л. Понятие лингвистической переменной и его применение к принятию приближенных решений. – М.: Мир, 1976. – 166 с.
22. Леоненков А. В. Нечеткое моделирование в среде MATLAB и fuzzyTECH. – СПб.: БХВ Петербург, 2005. – 736 с.: ил.
23. Malczewski J., 2006. A GIS-based multi-criteria decision analysis: A survey of the literature. *International Journal of Geographical Information Science* Vol. 20, No. 7, P. 703–726.
24. Одеса. Вікіпедія. [Электронный ресурс] – Режим доступа: <https://uk.wikipedia.org/wiki/Одеса>
25. Сервер OpenStreetMap. [Электронный ресурс] – Режим доступа: <https://www.openstreetmap.org>
26. GeoServer. [Электронный ресурс] – Режим доступа: <http://geoserver.org/>

27. PostGIS. [Электронный ресурс] – Режим доступа: <https://postgis>

28. OpenLayers. [Электронный ресурс] – Режим доступа: <http://openlayers.org/>

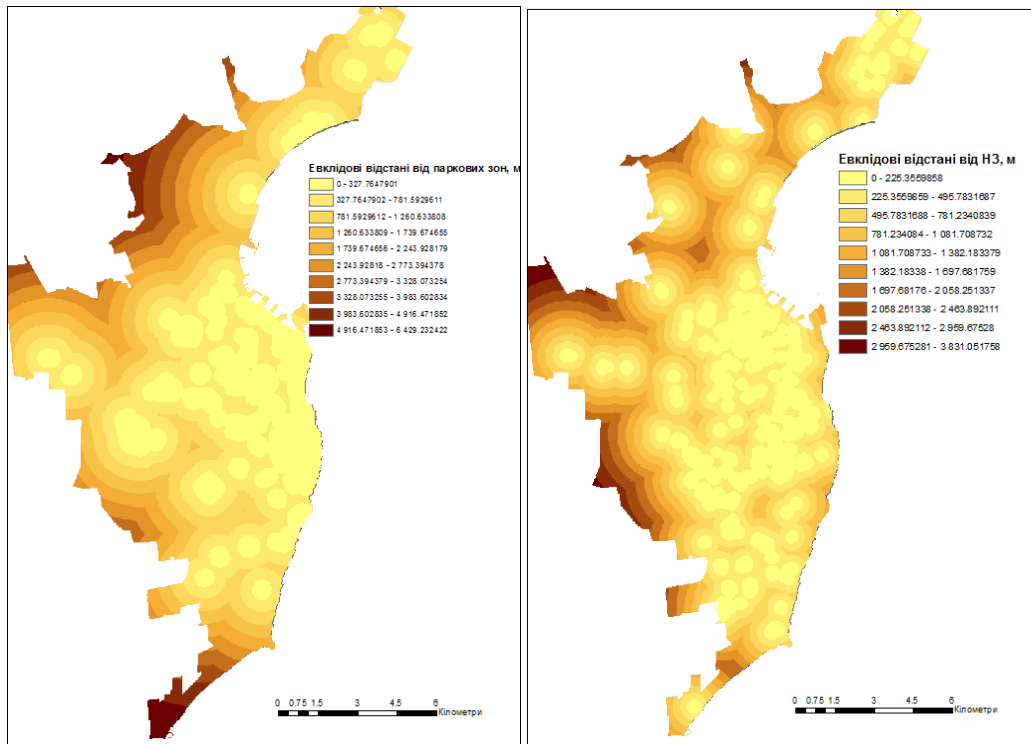
Д О Д А Т К И

ДОДАТОК А

Діаграма методології нечіткого зваженого накладання

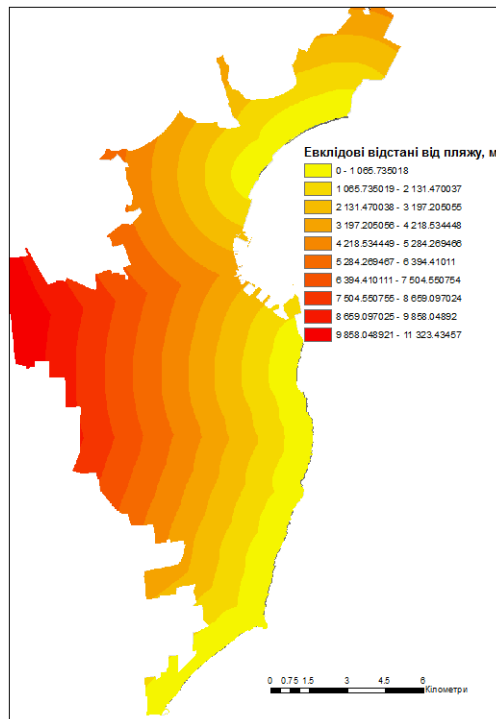


ДОДАТОК Б
Карти евклідових відстаней критеріїв



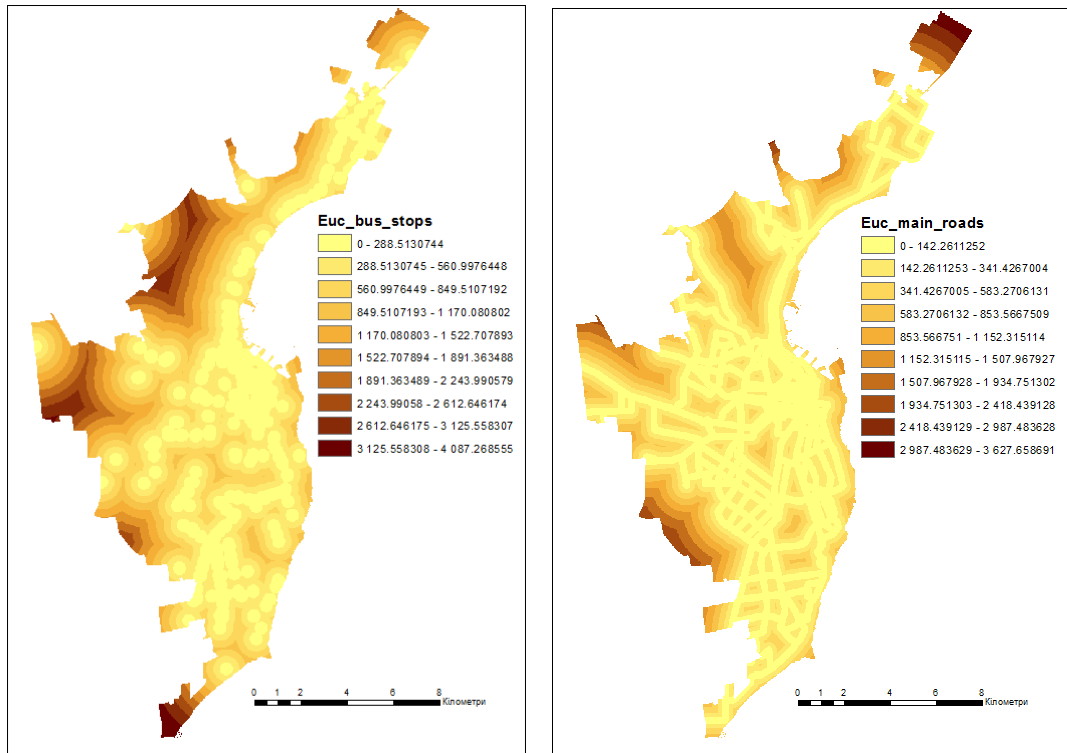
а

б



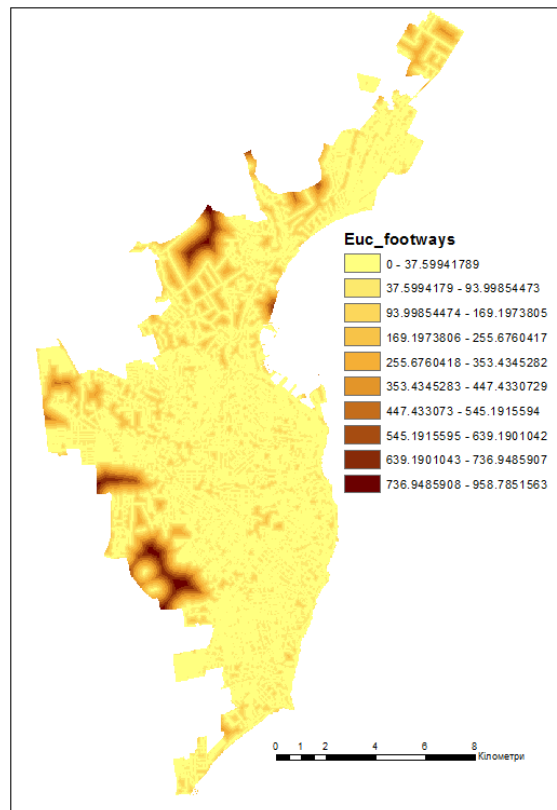
в

Рисунок Б.1 – Підкритерії «Наявність покупців»:
а – паркові зони; б – навчальні заклади; в – пляжні зони



а

б



в

Рисунок Б.2 – Підкритерії «Транспортна доступність»:
а – автобусні зупинки; б – головні дороги; в – пішохідні дороги

ДОДАТОК В

Матриця попарних порівнянь критеріїв

Критерії	Пляжні зони	Паркові зони	Навчальні заклади	Пішохідні дороги	Автобусні зупинки	Головні магістралі	Вага %
Пляжні зони	1	1	0,143	0,111	0,2	0,2	3,446
Паркові зони	1	1	0,143	0,143	0,2	0,2	3,612
Навчальні заклади	7	7	1	0,5	0,333	0,333	15,14
Пішохідні дороги	9	7	2	1	2	2	32,998
Автобусні зупинки	5	5	3	0,5	1	1	22,402
Головні магістралі	5	5	3	0,5	1	1	22,402

ДОДАТОК Г

Модель геобработки данных

