КИЇВСЬКИЙ НАЦІОНАЛЬНИЙ УНІВЕРСИТЕТ
БУДІВНИЦТВА І АРХІТЕКТУРИ

Факультет геоінформаційних систем та управління територіями
(факультет)

Кафедра геоінформатики і фотограмметрії
(назва кафедри)

ПОЯСНЮВАЛЬНА ЗАПИСКА
ДО АТЕСТАЦІЙНОЇ ВИПУСКНОЇ РОБОТИ
НА ЗДОБУТТЯ ОСВІТНЬОГО СТУПЕНЯ БАКАЛАВРА

на тему:

«Оцінювання якості відкритих геопросторових даних»

Божко Вероніка Костянтинівна
(прізвище, ім'я, по батькові студента повністю)

Київ 2025 р.

КИЇВСЬКИЙ НАЦІОНАЛЬНИЙ УНІВЕРСИТЕТ
БУДІВНИЦТВА І АРХІТЕКТУРИ

Факультет геоінформаційних систем та управління територіями
(факультет)

Кафедра геоінформатики і фотограмметрії
(назва кафедри)

ЗАТВЕРДЖУЮ
Завідувач кафедри
проф., д-р. техн. наук
Юрій КАРПІНСЬКИЙ
“____” _____2025 року

ПОЯСНЮВАЛЬНА ЗАПИСКА
ДО АТЕСТАЦІЙНОЇ ВИПУСКНОЇ РОБОТИ НА ЗДОБУТТЯ ОСВІТНЬОГО СТУПЕНЯ БАКАЛАВРА

«Оцінювання якості відкритих геопросторових даних»

Виконала: студентка групи ГІСТ-21
спеціальності 193
“Геодезія та землеустрій”
спеціалізації “Геоінформаційні системи і технології”
Божко Вероніка Костянтинівна
Керівник: Кінь Д.О., PhD, доцент
Ідентичність підтверджую

Київ 2025 р.

КИЇВСЬКИЙ НАЦІОНАЛЬНИЙ УНІВЕРСИТЕТ БУДІВНИЦТВА І АРХІТЕКТУРИ

Інститут, факультет Геоінформаційних систем і управління територіями
Кафедра Геоінформатики і фотограмметрії
Освітньо-кваліфікаційний рівень бакалавр
Напрям підготовки 19 «Архітектура та будівництво»
(шифр і назва)
Спеціальність 193 "Геодезія та землеустрій"
(шифр і назва)

ЗАТВЕРДЖУЮ
Завідувач кафедри
_____проф., д.т.н. Юрій КАРПІНСЬКИЙ
"__" _____2025 року

З А В Д А Н Н Я
НА ДИПЛОМНИЙ ПРОЕКТ (РОБОТУ) СТУДЕНТУ

Божко Вероніка Костянтинівна
(прізвище, ім'я, по батькові)

1. Тема проекту (роботи): «Оцінювання якості відкритих геопросторових даних»
керівник проекту (роботи) доцент Кінь Данило Олексійович
(прізвище, ім'я, по батькові, науковий ступінь, вчене звання)
затверджені наказом вищого навчального закладу від "__"__2025 року №
2. Строк подання студентом проєкту (роботи)_____
3. Вихідні дані до проекту (роботи):
   1) Ортофотоплан масштабу 1:10000
   2) Мозайка космічних знімків Google Satellietes
   3) Мозайка космічних знімків ESRI
   4) Мозайка космічних знімків Bing Maps
   5) Геопросторові дані OpenStreetMap
   6) Геопросторові дані Microsoft Bing Building Footprints
   7) Топографічні плани масштабу 1:2000 для населених пунктів: Тульчин, Лисянка, Славута, Пісочин та Рогань
4. Зміст розрахунково-пояснювальної записки (перелік питань, які потрібно розробити)

ВСТУП
Розділ 1. Аналіз методів оцінки якості відкритих геопросторових даних
1.1 Аналіз нормативно-правового забезпечення щодо якості геопросторових даних
1.2 Аналіз вимог до якості відкритих геопросторових даних
1.3 Аналіз міжнародного та вітчизняного досвіду щодо оцінки якості геопросторових даних
1.4 Метадані вихідних геопросторових даних проекту
Розділ 2. Методика оцінки якості геопросторових даних
2.1 Опис методики оцінки якості геопросторових даних
2.2 Концептуальна модель бази геопросторових даних для оцінювання якості даних
2.3 Логічна модель бази геопросторових даних для оцінювання якості даних
2.4 Опрацювання вихідних геопросторових даних проекту
Розділ 3. Дослідна реалізації методики оцінки якості відкритих геопросторових даних
3.1 Реалізація бази геопросторових даних проекту у середовищі ArcGIS

3.2 Підготовка ГІС-проекту для оцінювання якості відкритих геопросторових даних
3.3 Виконання оцінки якості відкритих геопросторових даних у середовищі ArcGIS
3.4 Аналіз результатів оцінки якості відкритих геопросторових даних

6. Консультанти розділів проекту (роботи)

Розділ	Прізвище, ініціали та посада консультанта	Підпис, дата	
		завдання видав	завдання прийняв
1	Професор Карпінський Ю.О		
2	Професор Карпінський Ю.О		
3			

7. Дата видачі завдання_____

КАЛЕНДАРНИЙ ПЛАН

№ з/п	Назва етапів дипломного проекту (роботи)	Строк виконання етапів проекту (роботи)	Примітка
1	ВСТУП	27.02.2025	
2	Аналіз нормативно-правового забезпечення щодо якості геопросторових даних	30.02.2025	
3	Аналіз вимог до якості відкритих геопросторових даних	02.03.2025	
4	Аналіз міжнародного та вітчизняного досвіду щодо оцінки якості геопросторових даних	05.03.2025	
5	Метадані вихідних геопросторових даних проекту	10.03.2025	
6	Опис методики оцінки якості геопросторових даних	14.03.2025	
7	Концептуальна модель бази геопросторових даних для оцінювання якості даних	20.03.2025	
8	Логічна модель бази геопросторових даних для оцінювання якості даних	21.03.2025	
9	Опрацювання вихідних геопросторових даних проекту	03.04.2025	
10	Реалізація бази геопросторових даних проекту у середовищі ArcGIS	07.04.2025	
11	Підготовка ГІС-проекту для оцінювання якості відкритих геопросторових даних	14.04.2025	
12	Виконання оцінки якості відкритих геопросторових даних у середовищі ArcGIS	22.04.2025	
13	Аналіз результатів оцінки якості відкритих геопросторових даних	09.05.2025	

Студент _____ Вероніка БОЖКО
(підпис) (прізвище та ініціали)

Керівник проекту (роботи) _____ Данило КІНЬ
(підпис) (прізвище та ініціали)

ЗМІСТ

ВСТУП…….……6
РОЗДІЛ 1. АНАЛІЗ МЕТОДІВ ОЦІНКИ ЯКОСТІ ВІДКРИТИХ ГЕОПРОСТОРОВИХ ДАНИХ…….……..….8
1.1 Аналіз нормативно-правового забезпечення щодо якості геопросторових даних…….8
1.2 Аналіз вимог до якості відкритих геопросторових даних……..…10
1.3 Аналіз міжнародного та вітчизняного досвіду щодо оцінки якості геопросторових даних…….……..12
1.4 Метадані вихідних геопросторових даних проєкту……18
РОЗДІЛ 2. МЕТОДИКА ОЦІНКИ ЯКОСТІ ГЕОПРОСТОРОВИХ ДАНИХ…..25
2.1 Опис методики оцінки якості геопросторових даних…….25
2.2 Концептуальна модель бази геопросторових даних для оцінювання якості даних……31
2.3 Логічна модель бази геопросторових даних для оцінювання якості даних……...35
2.4 Опрацювання вихідних геопросторових даних проєкту…….37
РОЗДІЛ 3. ДОСЛІДНА РЕАЛІЗАЦІЯ МЕТОДИКИ ОЦІНКИ ЯКОСТІ ВІДКРИТИХ ГЕОПРОСТОРОВИХ ДАНИХ……40
3.1 Реалізація бази геопросторових даних проєкту у середовищі ArcGIS...40
3.2 Підготовка ГІС-проекту для оцінювання якості відкритих геопросторових даних……...42
3.3 Виконання оцінки якості відкритих геопросторових даних у середовищі ArcGIS…….…45
3.4 Аналіз результатів оцінки якості відкритих геопросторових даних..…51
ВИСНОВОК……...53
СПИСОК ВИКОРИСТАНИХ ДЖЕРЕЛ……..54

ВСТУП

У XXI столітті геопросторові дані перетворилися на один із ключових інформаційних ресурсів, що сприяють ухваленню рішень у широкому спектрі напрямків – від державного управління, містобудування та сільського господарства до екологічного моніторингу, забезпечення безпеки й реагування на надзвичайні ситуації. З поступовим поширенням ідеї відкритих даних державні органи все частіше надають суспільству доступ до наборів геоінформації. Це відкриває нові перспективи для науки, бізнесу та громадськості. Однак при цьому виникає важливе питання: наскільки якісними та придатними для практичного використання є ці відкриті дані?

Якість геопросторових даних має фундаментальне значення для точності аналітичних розрахунків і результативності ГІС-рішень. Дані низької якості можуть стати причиною помилкових управлінських рішень, неправильного оцінювання ризиків і некоректного аналізу ситуацій. Тому важливо застосовувати формалізовані підходи до перевірки якості відкритих геопросторових даних, керуючись міжнародними стандартами (ISO 19157, INSPIRE) та адаптуючи їх до вітчизняного правового середовища.

Ця робота зосереджена на дослідженні проблемного питання якості відкритих геопросторових даних для використання їх у топографо-геодезичній та картографічній діяльності, національної інфраструктури геопросторових даних, містобудівної та землевпорядної діяльностях. Пояснювальна записка містить три основні частини. Перший розділ включає аналіз нормативно-правової бази й міжнародного досвіду, визначення вимог до якості та опис метаданих наборів даних, які були використані у дослідженні. У другому розділі запропоновано методику оцінювання якості геопросторових даних, побудовано концептуальну і логічну моделі бази даних для відображення показників якості, а також виконано підготовку даних до подальшого аналізу. У третьому розділі реалізовано зазначену методику в програмно-аналітичному середовищі ArcGIS: створено ГІС-проєкт, виконано оцінювання і аналіз отриманих результатів.

Метою роботи є виконання оцінки якості відкритих геопросторових даних із використанням сучасних геоінформаційних технологій. Досягнення цієї мети сприятиме ефективнішому застосуванню геопросторових даних завдяки забезпеченню їхньої надійності та відповідності потребам користувачів для прийняття обґрунтованих рішень у різних сферах діяльності.

Основні завдання:

1) проаналізувати нормативно-правове забезпечення щодо якості геопросторових даних та вимог до неї;

2) дослідити міжнародний та вітчизняний досвід щодо виконання оцінки якості геопросторових даних;

3) сформувати метадані вихідних наборів геопросторових даних проєкту та опрацювати їх;

4) описати методику оцінки якості геопросторових даних;

5) розробити концептуальну модель бази геопросторових даних для оцінювання якості даних;

6) розробити логічну модель зазначеної бази геопросторових даних;

7) реалізувати базу геопросторових даних проєкту у середовищі ArcGIS;

8) виконати оцінку якості наборів відкритих геопросторових даних у середовищі ArcGIS;

9) проаналізувати отримані результати та сформувати висновки та рекомендації щодо використання наборів відкритих геопросторових даних.

Об’єкт дослідження: набори відкритих геопросторових даних.

Предмет дослідження: якість наборів відкритих геопросторових даних.

РОЗДІЛ 1. АНАЛІЗ МЕТОДІВ ОЦІНКИ ЯКОСТІ ВІДКРИТИХ ГЕОПРОСТОРОВИХ ДАНИХ

1.1 Аналіз нормативно-правового забезпечення щодо якості геопросторових даних

Аналіз нормативно-правового забезпечення якості геопросторових даних в Україні враховує міжнародні стандарти, чинне національне законодавство та підзаконні акти, що регулюють процеси створення, обробки, зберігання і застосування геопросторової інформації.

1. Закон України «Про національну інфраструктуру геопросторових даних» (№ 554-IX від 13.04.2020). Цей Закон слугує основою для формування ключових термінів і понять, необхідних у дипломній роботі. Він регулює правові аспекти створення та функціонування Національної інфраструктури геопросторових даних (НІГД). Положення документа застосовуються для визначення стандартів якості геоданих, забезпечення їхньої доступності, а також окреслення обов'язків власників даних. Це дає змогу проаналізувати, наскільки структура досліджуваних даних відповідає вимогам чинного законодавства [1].

2. Технічні вимоги та методи забезпечення інтероперабельності і сумісності наборів геопросторових даних та геоінформаційних сервісів (Наказ Мінагрополітики № 347 від 10.11.2021). У практичному розділі дипломної роботи ці вимоги використовуються для розробки критеріїв оцінювання сумісності геопросторових сервісів і форматів даних. Розглядається відповідність наданої інформації визначеній еталонній моделі, а також здатність забезпечувати базову міжвідомчу взаємодію [2].

3. Технічні вимоги до метаданих для наборів геопросторових даних і геоінформаційних сервісів національної інфраструктури геопросторових даних (Наказ Мінагрополітики № 347 від 10.11.2021). Зазначений документ слугує для оцінки наявності, повноти та структури метаданих, які супроводжують геопросторові дані. У рамках дипломної роботи було здійснено аналіз конкретного набору відкритих даних на відповідність цим вимогам [3].

4. Про затвердження Порядку функціонування національної інфраструктури геопросторових даних (№ 532-2021-п; Постанова Кабінету Міністрів України; Порядок, Форма типового документа, Заява від 26.05.2021 №532). Цей документ слугує для опису життєвого циклу геопросторових даних – від їх створення до оновлення та розповсюдження. У ньому розглядається практична реалізація цього циклу на прикладі наборів даних, які доступні на офіційних державних порталах [4].

5. Директива 2007/2/ЄС про створення Інфраструктури просторової інформації у Європейському Співтоваристві (INSPIRE). Директива слугує міжнародним орієнтиром для порівняння української НІГД з європейською моделлю. У дипломній роботі на її основі проводиться порівняльний аналіз структури, відкритості та якості метаданих, а також визначаються можливі шляхи адаптації до європейських стандартів [5].

6. ISO 19100 (серія міжнародних стандартів для геоінформації). Ключовим стандартом є ISO 19157:2013 «Географічна інформація – Якість даних». Цей міжнародний стандарт використовується для розроблення методології перевірки якості геопросторових даних. У ньому визначено основні характеристики якості, такі як повнота, точність, логічна узгодженість та актуальність, які були залучені під час аналізу відкритих геоданих у практичній частині роботи [6].

7. Open Geospatial Consortium (OGC) Standards. Ці стандарти, зокрема WMS та WFS, розглядаються у контексті оцінки інтероперабельності сервісів, що забезпечують доступ до геоданих. У дипломній роботі аналізується відповідність досліджуваних ресурсів основним протоколам обміну просторовою інформацією [8].

1.2 Аналіз вимог до якості відкритих геопросторових даних

Відповідно до Закону України «Про національну інфраструктуру геопросторових даних» базові геопросторові дані – загальнодоступні геопросторові дані, що складають уніфіковану цифрову координатно-просторову основу для виробництва, інтеграції та провадження іншої діяльності з різними геопросторовими даними [1].

Базовий набір геопросторових даних включає: системи відліку координат і висот; державний кордон України; адміністративно-територіальні одиниці разом із їхніми межами; територіальні громади, включаючи межі їхніх територій; гідрографічні об’єкти та гідротехнічні споруди; населені пункти з їхньою вулично-дорожньою мережею; будівлі та споруди; автомобільні дороги; залізничні шляхи; інженерні комунікації; аеропорти, морські та річкові порти; земний покрив і ґрунти; земельні ділянки; реєстри вулиць і адреси об’єктів; географічні назви; цифрову модель рельєфу; ортофотоплани. Ці дані не повинні містити інформації, що становить державну таємницю або доступ до якої обмежено відповідно до закону. Тематичні геопросторові дані формуються на основі базових геопросторових даних.

Процес створення геопросторових даних передбачає збір даних за допомогою наземних вимірювань, супутникових знімків та автоматизованих систем моніторингу. Зібрана інформація проходить обробку, що включає стандартизацію форматів, перевірку на точність і виправлення можливих помилок. Після цього дані публікуються у відкритих реєстрах, на державних геопорталах або міжнародних платформах, а також регулярно оновлюються для забезпечення їх актуальності.

Геопросторові дані є відкритими, оскільки це сприяє забезпеченню прозорості державного управління, підтримує розвиток бізнесу, наукових досліджень та громадських ініціатив. Вільний доступ до картографічної інформації має ключове значення для підвищення безпеки, ефективного містобудування, екологічного моніторингу, транспортного планування й

реагування у надзвичайних ситуаціях. В Україні відкритість геоданих регулюється Законом «Про національну інфраструктуру геопросторових даних», який забезпечує всім користувачам безкоштовний доступ до цієї інформації [1].

Одним із найважливіших аспектів якості геопросторових даних є дотримання вимог до точності топографічних планів та цифрових моделей ситуації. Згідно з пунктами 24 – 28 «Порядку створення топографічних планів та цифрових моделей» [13], затвердженого постановою Кабінету Міністрів України, визначено такі нормативи:

Планова точність: середні похибки положення об’єктів із чіткими обрисами відносно найближчих точок геодезичної основи не повинні перевищувати 0,5 мм у масштабі плану (наприклад, 1 м у масштабі 1:2000). У гірських і лісових районах ця похибка може становити до 0,7 мм, а на територіях з щільною капітальною забудовою – не більше 0,4 мм між об’єктами.

Допустимість зниження точності: відповідно до технічного завдання допускається створення планів зі зниженою графічною точністю – наприклад, у масштабі 1:2000 із точністю плану масштабу 1:5000. У таких випадках на плані обов’язково зазначається метод створення (зйомка, фотоплани, масштабування тощо) та точність виконання.

Висотна точність: середньоквадратичні похибки висот у цифровій моделі рельєфу не повинні перевищувати значень, встановлених у відповідних таблицях нормативного документа (залежно від масштабу зйомки).

Контроль якості: точність планів визначається шляхом порівняння положення контурів і висот точок, отриманих за контрольними вимірюваннями. Граничні розходження не повинні перевищувати подвоєні значення допустимих похибок і можуть стосуватись не більше ніж 10% вимірювань. Окремі перевищення (до 5% від вимірів) допускаються за умови подвоєної похибки.

Роздільна здатність векторних моделей: максимально допустима відстань між вершинами контурів векторних об’єктів залежить від масштабу і становить:

– 60 м для масштабу 1:5000;

– 40 м для масштабу 1:2000;

– 20 м для масштабу 1:1000;

– 15 м для масштабу 1:500.

Виконання зазначених вимог є важливою умовою забезпечення високої якості цифрових геопросторових моделей та їхньої придатності до використання в кадастрових, інженерних та планувальних задачах. У процесі оцінювання відкритих геоданих відповідність цим критеріям виступає основним орієнтиром для визначення точності та детальності вхідних матеріалів.

Проблема якості геопросторових даних стає особливо важливою, коли вони використовуються в геоінформаційних системах для запобігання та оперативного реагування на надзвичайні ситуації. У таких випадках від точності, повноти та актуальності цих даних залежить не лише масштаб матеріальних збитків, але й нерідко життя людей. Кожна сфера діяльності висуває унікальні вимоги до структури та змісту геоінформаційних моделей, включаючи набір об’єктів і їхні атрибути, точність просторових характеристик, а також топологічні й логічні зв’язки між об’єктами, які повинні бути коректно відображеними та реалізованими. Окремої уваги заслуговує створення цифрових моделей місцевості, яке є не лише складним, але й затратним процесом. З огляду на це, необхідно розробляти способи формування багатоцільових моделей місцевості, що дозволять забезпечити міжгалузеву інтеграцію та інтероперабельність геоінформаційних даних. Важливо також визначати загальні принципи й логічні підходи до побудови прикладних моделей, а головне – створити концептуальні основи, методи оцінки якості та критерії для їхнього удосконалення.

1.3 Аналіз міжнародного та вітчизняного досвіду щодо оцінки якості геопросторових даних

Оцінка якості геопросторових даних є важливим елементом ефективного управління просторовою інформацією, що відображено як у міжнародних нормативних актах, так і в законодавстві України. Однією з основ сучасного розуміння якості геопросторових даних є Директива INSPIRE (2007/2/ЄС) та

міжнародний стандарт ISO 19157:2013 [5,6]. Ці документи визначають не лише основні показники якості, але й підходи до її вимірювання, аналізу та забезпечення на всіх етапах життєвого циклу геоданих.

У Директиві INSPIRE якість геопросторових даних розглядається як багатогранна характеристика, яка охоплює різні аспекти: від просторової точності до логічної узгодженості та актуальності. Документ зобов'язує країни-члени ЄС забезпечити наявність метаданих, які описують якість даних, методи їх збору, часові рамки, обмеження у використанні, джерела та рівень точності. Оцінювання якості даних у межах INSPIRE враховує сферу їх застосування. Наприклад, для екологічного моніторингу пріоритетною є тематична точність, тоді як для територіального планування критично важлива просторово-геометрична точність. INSPIRE підтримує повторне використання геопросторових даних, тому особливий акцент робиться на прозорості та стандартизованому описі параметрів якості через метадані. Це дозволяє користувачам приймати обґрунтовані рішення щодо придатності даних для конкретних завдань.

Міжнародний стандарт ISO 19157 деталізує підхід до контролю якості геоданих і є основоположним технічним документом у цій сфері [6]. У стандарті визначено чіткі категорії якості: повнота (наскільки повні дані, чи є прогалини), логічна послідовність (внутрішня узгодженість даних), просторово-часова точність (точність координат та часових показників), тематична правильність (відповідність даних реальним характеристикам об'єктів) і правильність атрибутів (точність описових характеристик об'єктів). Для кожної категорії передбачено відповідні міри – наприклад, середня квадратична похибка, відсоток помилок або ступінь відхилення. ISO 19157 також передбачає ведення формалізованих звітів про якість, що повинні бути пов'язані з метаданими (відповідно до ISO 19115), дозволяючи не лише оцінити якість, а й відслідковувати її зміну в процесі оновлення або трансформації даних [6].

Відповідно до ISO 8402, якість визначається як сукупність характеристик продукту, що впливають на його здатність задовольняти встановлені чи

передбачувані потреби [9]. В ISO 15046, єдиному стандарті ISO, присвяченому геоінформації, який ще перебуває в розробці технічним комітетом (TC) 211 [10], просторову якість визначають як одну чи кілька характеристик географічних даних, що описують ступінь їхньої придатності для використання.

Узявши до уваги всі ці визначення, просторову якість даних можна визначити як сукупність показників, які відображають ствердні чи заперечні відмінні властивості та характеристики географічних даних. Вони описують відповідність специфікаціям та придатність для використання з точки зору як виробників, так і користувачів. Із цього визначення видно, що просторова якість є суб'єктивною характеристикою, яка по-різному сприймається користувачами й виробниками. Наприклад, набір просторових даних може відповідати специфікаціям виробника і тому бути оціненим як якісний, але водночас не відповідати очікуванням іншого користувача чи виробника. Аналогічно, для користувачів такий набір даних може бути або відповідним до їхніх сфер застосування, або ні.

Три основні фактори, що визначають просторову якість даних:

– кошти (витрати для виробників, ціна для користувачів);

– час (період виробництва для виробників, час доступу до даних для користувачів);

– прийнятний рівень якості (мети виробництва для виробників, сфери застосування для користувачів).

Ці фактори є багатогранними. Виробники можуть знижувати якість задля зменшення витрат і скорочення часу виробництва, тоді як користувачі прагнуть отримати дані за низькою ціною й за короткий проміжок часу, при цьому готові миритися з певним рівнем зниження якості. Якщо користувачі хочуть підвищити якість даних, вони мають бути готовими платити більше та довше чекати на отримання інформації. Щодо прийнятного рівня якості він може різнитися між виробниками та користувачами залежно від виробничих цілей перших і сфер застосування останніх.

У вітчизняній практиці оцінювання якості геопросторових даних одним із найпоширеніших джерел перевірки є ортофотоплани масштабу 1:10 000, створені у межах реалізації проєкту Світового банку з оновлення топографічної основи України. Ці матеріали були сформовані на основі аерофотозйомки, виконаної у період 2006–2012 років, та опрацьовані з використанням технології орторектифікації, що забезпечує геометричну відповідність зображених об’єктів. Ортофотоплани мають просторову роздільну здатність 0,6 метра на піксель, що дозволяє розрізняти об’єкти інфраструктури, контури будівель і транспортні мережі.

Згідно зі сталою практикою в Україні, ці ортофотоплани активно використовуються як еталон для перевірки векторизованих об’єктів, позиційної точності кадастрових даних, а також для візуального аналізу відповідності фактичній ситуації. Вони є набором геопросторових даних, який можна віднести до загальнодержавного рівня.

До основних переваг використання ортофотопланів 1:10 000 віднесено:

– висока деталізація та роздільна здатність, яких достатньо для ідентифікації об’єктів місцевості;

– виконання геометричної корекції (орторектифікації);

– охоплення великої території та стандартизований підхід до опрацювання;

– офіційне джерело походження, що підвищує довіру до їхньої якості.

Серед недоліків варто зазначити:

– застарілість даних (відсутність оновлень після 2012 року), що знижує їхню актуальність для сучасної забудови;

– обмежену позиційну точність (декларовану як 5–7 м), що ускладнює виконання завдань, де потрібна точність на рівні 1 м або вище;

– не суцільне покриття зображення, оскільки виявлені .

Таким чином, досвід використання ортофотопланів у вітчизняних проєктах демонструє прагнення до об’єктивного оцінювання геопросторових даних навіть у випадку обмеженої доступності актуальних джерел.

Міжнародна практика оцінювання якості геопросторових даних представлена широким спектром підходів, які базуються на стандартах ISO, INSPIRE, OGC та практичних прикладах впровадження в окремих країнах. Одним із ключових джерел є стаття Jingfeng Xia «Metrics to Measure Open Geospatial Data Quality» (2012) [14], де досліджено принципи розробки метрик для оцінки відкритих геоданих. У роботі запропоновано концептуальну модель, що поєднує як кількісні, так і якісні показники. Серед основних критеріїв автор виділяє точність, повноту, сумісність, актуальність, легкість використання, а також прозорість та доступність метаданих. Значна увага приділяється інтероперабельності – здатності даних бути використаними в різних інформаційних системах, що є важливим фактором при інтеграції відкритих джерел у ГІС-проєкти. Цей підхід був врахований у дипломній роботі для формування власної системи оцінювання даних, яка охоплює як просторові (позиційна точність), так і описові (детальність, повнота) характеристики.

Важливий теоретичний і методичний внесок представлений у статті Hayati Taştan та M. Orhan Altan «Spatial Data Quality» [15], де розглядаються джерела просторових похибок та систематизуються методи оцінювання якості. Автори класифікують якість за такими складовими: позиційна точність, логічна узгодженість, повнота, тематична точність і актуальність. Для перевірки якості пропонується три методи: повна інспекція, вибіркове тестування та непряма оцінка. Основна ідея полягає у тому, що якість має контролюватися на всіх етапах життєвого циклу даних, із врахуванням витрат і потреб користувача. У дипломній роботі ці методичні підходи були використані при побудові практичної оцінки: здійснено вибіркову перевірку контрольних будівель, використано обґрунтовані порогові значення довжин сторін полігонів, реалізовано багатоступеневу перевірку точності.

Практичне дослідження, що оцінює точність цифрових моделей висот, наведене у статті Dawod & Ascoura «The Validity of Open Source Elevations for National-Scale Flood Inundation Modeling» (2021) [16]. У ній автори порівнюють DEM з відкритих джерел (SRTM, ASTER, ALOS) з контрольними даними в межах

Саудівської Аравії. Застосовано метод середньоквадратичної похибки (RMSE), а також аналіз просторової варіативності похибок у гірських і рівнинних районах. Встановлено, що похибки можуть досягати 8 метрів, що неприйнятно для високоточного картографування масштабу 1:25 000 і менше. Незважаючи на те, що дипломна робота орієнтується на векторні дані, принцип застосування RMSE як кількісної метрики точності був адаптований для оцінки позиційного відхилення точок та геометричних параметрів будівель у межах еталонних полігонів.

Оцінювання добровільних геоданих розглядається у статті Moradi et al. «Evaluating OSM Building Footprint Data Quality in Québec Province, Canada» (2023) [17]. Дослідження базується на порівнянні контурів будівель OpenStreetMap із офіційними муніципальними джерелами для п’яти міст Канади. Проаналізовано чотири параметри якості: повнота (відношення кількості OSM-об’єктів до еталонних), точність позиціонування, геометрична відповідність (форма будівлі) та атрибутивна коректність. Результати показали значне зростання якості OSM у часовому аспекті (2018–2023), особливо в активних громадах. У дипломній роботі ці підходи були враховані для побудови таблиць повноти та детальності, які базуються на відсотковому співвідношенні об’єктів, що задовольняють певним критеріям. Також використано ідею багаторівневої класифікації: повна відповідність, часткова відповідність, невідповідність.

Таким чином, міжнародні дослідження підтверджують необхідність комплексного оцінювання якості відкритих геопросторових даних, із застосуванням формалізованих критеріїв і адаптацією до конкретного випадку використання. Отримані в результаті аналізу підходи, методи та метрики були адаптовані до українських умов і враховані при розробці власної методики оцінювання, реалізованої в середовищі ArcGIS у межах дипломного дослідження.

1.4 Метадані вихідних геопросторових даних проєкту

Для цього дослідження були використані геопросторові дані з відкритого доступу, надані через офіційні державні геоінформаційні сервіси та міжнародні публічні платформи. Усі джерела попередньо оцінювались на наявність базових метаданих, які містять інформацію про формат, систему координат, дату оновлення, доступність, рівень деталізації та обмеження у використанні. Оформлення метаданих наборів геопросторових даних та геоінформаційних сервісів (табл. 1.1) здійснювалося відповідно до Технічних вимог до метаданих НІГД (наказ Мінагрополітики № 347 від 10.11.2021) та міжнародним стандартом ISO 19115 [3,7]. Як приклад, у межах дослідження було оформлено таблицю метаданих лише для набору геопросторових даних OpenStreetMap.

У рамках проєкту було обрано п’ять основних джерел геопросторових даних, які стали об’єктами оцінки якості:

1) OpenStreetMap (OSM) – векторна база даних, створена спільнотою користувачів. Інформація про будівлі експортується за допомогою плагіна QuickOSM у форматі .osm і конвертується у .shp. Джерело надає структуровані геометричні й атрибутивні дані, проте може мати обмеження щодо повноти і рівномірності точності для всіх територій.

2) Microsoft Bing Building Footprints – набір векторних полігонів будівель, згенерованих автоматично з використанням нейромереж. Дані завантажуються з відкритого репозиторію GitHub у форматі GeoJSON. Хоч цей ресурс і не містить метаданих для кожного об’єкта, він охоплює значну територію з високим рівнем деталізації.

3) Google Satellite – супутникова мозаїка високої роздільної здатності, інтегрована в QGIS через XYZ Tile Layer із використанням шаблону: http://mt0.google.com/vt/lyrs=s&hl=en&x={x}&y={y}&z={z} . Ця мозаїка є виключно візуальною підкладкою для ручного оцифровування контурів будівель. Векторні об’єкти створюються користувачем вручну на основі зображень.

4) Bing Maps – супутникові знімки Microsoft, підключені як растрова основа через шаблон XYZ: http://ecn.t3.tiles.virtualearth.net/tiles/a{q}.jpeg?g=1 Використовується аналогічно до Google Satellite, лише для візуального ідентифікування об'єктів під час векторизації.

5) Esri Imagery – супутниковий сервіс від компанії Esri, підключається у вигляді XYZ Tile Layer через URL-шаблон: https://server.arcgisonline.com/arcgis/rest/services/World_Imagery/MapServer/tile/{z}/{y}/{x}. Ці дані є фоновим зображенням для візуального позиціонування та оцифрування об'єктів. Сервіс пропонує високу деталізацію й регулярно оновлюється. Векторні контури будівель створювались вручну на основі супутникових знімків.

Таблиця 1.1

Метадані набору відкритих геопросторових даних OpenStreetMap

№ з/п	Назва елемента метаданих та його познака за ISO 19115-1:2014	Коментар	Значення метаданих
1	2	3	4
1	Унікальний ідентифікатор метаданих: (MD_Metadata.metadataIdentifier: MD_Identifier)	Унікальний ідентифікатор для метаданих, що формується програмними засобами підготовки метаданих	–
2	Назва ресурсу: (MD_Metadata.identificationInfo > MD_DataIdentification.citation > CI_Citation.title)	Характерна унікальна назва, що ідентифікує ресурс (набір даних або комплект наборів даних)	OpenStreetMap
3	Дата створення/оновлення ресурсу: (MD_Metadata.idenitificationInfo > MD_DataIdentification.citation > CI_Citation.date та CI_Citation.dateType)	за класифікатором CI_DateTypeCode	
		створення	2004-08-09T09:00:00
		оновлення	2025-02-09T21:21:00
4	Контактна інформація відповідальних за ресурс: (MD_Metadata.identificationInfo >	за класифікатором CI_RoleCode	
		постачальник	OpenStreetMap Foundation
		держатель	OpenStreetMap Foundation

№ з/п	Назва елемента метаданих та його познака за ISO 19115-1:2014	Коментар	Значення метаданих
	MD_DataIdentification.PointOfContact > CI_Responsibility)		
5	Просторове охоплення: (MD_Metadata.identificationInfo > MD_DataIdentification.extent > EX_Extent.geographicElement > EX_GeographicExtent > EX_GeographicBoundingBox або EX_BoundingPolygon або EX_GeographicDescription)	Опис просторового охоплення території даними ресурсу	EPSG:3857 - 20037508.3427892439067364,-20037508.3427892476320267 : 20037508.3427892439067364,20037508.3427892476320267
6	Мова ресурсу: (MD_Metadata > MD_DataIdentification.defaultLocale > PT_Locale)	Код мови та код набору символів, що використовуються в ресурсі. За замовчуванням: код мови: Ukr; код набору символів: utf-8.	UTF-8
7	Тематична категорія ресурсу: (MD_Metadata.identificationInfo > MD_DataIdentification.topicCategory > MD_TopicCategoryCode)	за класифікатором базових категорій, що визначені в ISO 19115-1:2014, який може доповнюватися.	Межі; Рельєф; Базові зображення поверхні Землі;
8	Просторове розрізнення: (MD_Metadata.identificationInfo > MD_Identification.spatialResolution > MD_Resolution.equivalentScale, MD_Resolution.distance, MD_Resolution.vertical або MD_Resolution.angularDistance або MD_Resolution.levelOfDetail)	Числовий масштаб (рівень деталізації)	Залежить від джерел даних, території
9	Тип ресурсу: (MD_Metadata.metadataScope > MD_Scope.resourceScope)	за класифікатором MD_ScopeCode	dataset
10	Стислий опис ресурсу: (MD_Metadata.identificationIn	Реферативний опис, призначення, структури, складу та змісту набору	OpenStreetMap — це глобальна вільна картографічна платформа, створена для збору, збереження

№ з/п	Назва елемента метаданих та його познака за ISO 19115-1:2014	Коментар	Значення метаданих
	fo > MD_Identification.abstract)	геопросторових даних та інші відомості про ресурс.	та поширення геопросторових даних, доступних для всіх. Вона призначена для використання в навігаційних системах, геоінформаційних системах (ГІС), містобудівній діяльності, наукових дослідженнях, веб-картах і мобільних додатках. Склад та зміст набору: адміністративні межі (країни, області, громади); транспортна інфраструктура (вулиці, залізниці, мости, тунелі); гідрографія (річки, озера) житлова та комерційна забудова зелені зони та землекористування інфраструктурні об'єкти (школи, лікарні, АЗС, магазини) мобільна інфраструктура (велодоріжки, зупинки, маршрути); історичні, культурні та туристичні об'єкти
11	Додаткова інформація про часове та/або вертикальне охоплення ресурсу: (MD_Metadata.identificationInfo > MD_Identification.extent > EX_Extent > EX_TemporalExtent або EX_VerticalExlent)	Вказується період часу, охоплений вмістом ресурсу (тип даних TM_Primitive за ДСТУ ISO 19108:2017) та/або інтервал значень відміток висоти у визначеній системі висот.	to 2025-02-09T21:21:00
12	Походження ресурсу: (MD_Metadata > resourceLineage > LI_Lineage)	Опис джерел ресурсу та основних технологічних процесів, що використовувалися для виробництва ресурсу.	Багато даних в OSM вручну векторизовані з аерофотозйомки чи супутникових зображень високої роздільної здатності (наприклад, Bing, Maxar). Часто прив'язані до GPS-треків, що забезпечує високу точність. У містах рівень деталізації іноді перевищує 1 м (особливо, якщо активно картографують місцеві волонтери).
13	Онлайн доступ до ресурсу: (MD_Metadata.identificationInfo >	Вказуються URL адреси онлайн доступу до ресурсу,	https://www.openstreetmap.org/

№ з/п	Назва елемента метаданих та його познака за ISO 19115-1:2014	Коментар	Значення метаданих
	MD_DataIdentification.citation > CI_Citation.onlineResource > CI_OnlineResource)	геоінформаційних сервісів та додаткової інформації про ресурс.	
14	Ключові слова: (MD_Metadata > MD_DataIdentification > descriptiveKeywords > MD_Keywords)	Слова або фрази, що описують ресурс та використовуються для індексування в пошукових системах.	Топографія, карта, ГІС, відкриті дані
15	Обмеження щодо доступу та використання ресурсу: (MD_Metadata > MD_DataIdentification > MD_Constraints > MD_Constraints.useLimitations та/або MD_LegalConstraints та/або MD_SecurityConstraints)	за класифікатором MD_RestrictionCode	Дані доступні під ліцензією Open Database License (ODbL). Вимагається атрибуція «© OpenStreetMap contributors
16	Дата створення/оновлення метаданих: (MD_Metadata.dateInfo:CI_Date)	Дата створення та оновлення метаданих	14.05.2025
17	Контактна інформація відповідальних за метадані: (MD_Metadata.contact: CI_Responsibility)	Назва та контактні дані сторони, відповідальної за створення та оновлення метаданих	Вероніка Божко, студентка 4-ого курсу групи ГІСТ-21, КНУБА
18	Інформація про референцну систему координат (SRC) ресурсу: (MD_Metadata > MD_ReferenceSystem > referenceSystemIdentifier: MD_Identifier)	Ідентифікатор (код) SRC та системи висот згідно з базою даних міжнародного репозитарію параметрів систем координат EPSG.	EPSG: 3857
19	Статус ресурсу: (MD_Metadata.identificationInfo > MD_Identification.status > MD_ProgressCode)	за класифікатором MD_ProgressCode	Поточний, дійсний
20	Інформація про відповідність ресурсу: (MD_DataQualityCommon > DQ_Result > DQ_ConformanceResult > *обов'язково* DQ_ConformanceResult.specification: CI_Citation та DQ_ConformanceResult. pass: Boolean і за потреби	Подається як клас метаданих про якість даних у відповідності до ISO 19157:2013. Обов'язково вказується посилання на специфікацію набору геопросторових даних або на інший документ з вимогами до даних та	OpenStreetMap Data Manual https://wiki.openstreetmap.org/wiki/Good_practice Інші джерела специфікацій: Tagging Guidelines ODbL License Overpass API Data Model

№ з/п	Назва елемента метаданих та його познака за ISO 19115-1:2014	Коментар	Значення метаданих
	DQ_ConformanceResult.Explanation: CharacterString)	показник відповідності (1= так; 0= ні), а за потреби надається текстове пояснення рівня відповідності	
21	Інформація про формат даних ресурсу: (MD_Metadata > MD_Format.formatSpecificationCitation: CI_Citation)	Вказується загальноприйнята познака формату даних, версії формату та/або URL адреса посилання на специфікацію для формату даних	osm (XML) pbf (бінарний формат) GeoJSON, KML, Shapefile (через перетворення) PostgreSQL/PostGIS (через імпорт) Доступні експорти у форматі .PBF або GeoJSON.
22	Інформація про підтримку ресурсу: (MD_Metadata > MD_MaintenanceInformation > MD_MaintenanceInformation.maintenanceAndUpdateFrequency або userDefinedMaintenanceFrequency)	за класифікатором MD_Maintenance FrequencyCode або визначена виробником періодичність оновлення	Дані оновлюються в режимі реального часу (залежно від території)
23	Інформація про мову метаданих: (MD_Metadata.defaultLocale: PT_Locale)	Код мови та код набору символів, що використовуються в ресурсі. За замовчуванням: код мови: Ukr; код набору символів: utf-8	UTF-8
24	Інформація про відповідність метаданих стандартам: (MD_Metadata.metadataStandard: CI_Citation та/або MD_Metadata.metadataProfile: CI_Citation)	Вказується позначення та назва стандарту (стандартів) та/або профілю, яким відповідає набір елементів метаданих	ISO 19115-1:2014

Висновок до розділу 1: У результаті проведеного аналізу встановлено, що якість геопросторових даних є критично важливим фактором для їх ефективного застосування в управлінні територіями, містобудуванні, екологічному моніторингу та інших сферах. Вивчення нормативно-правового забезпечення засвідчило наявність достатньої законодавчої бази в Україні, яка узгоджена з міжнародними стандартами, такими як INSPIRE та ISO 19157. Систематизовано основні вимоги щодо структури, точності, повноти та метаданих відкритих геопросторових джерел. Проведено також огляд практичного досвіду оцінки якості даних, який підтвердив необхідність формалізованого підходу до перевірки відповідності даних їхньому призначенню.

РОЗДІЛ 2. МЕТОДИКА ОЦІНКИ ЯКОСТІ ГЕОПРОСТОРОВИХ ДАНИХ

2.1 Опис методики оцінки якості геопросторових даних

Використовуючи специфікацію продукту, виробник даних визначає елемент або елементи якості даних, які потрібно оцінити для перевірки відповідності специфікації. Зі свого боку, користувач даних, орієнтуючись на свої вимоги, обирає елемент чи елементи якості даних, які мають бути проаналізовані для оцінки того, наскільки набір даних відповідає його потребам. Елементи якості даних включають такі категорії, як повнота, логічна узгодженість, позиційна точність, часову точність і тематичну точність.

Відповідно до специфікації продукту, виробник визначає піделемент або піделементи якості даних, які необхідно оцінити для перевірки з дотриманням специфікації. Водночас користувач даних обирає піделементи елементів якості даних на основі своїх вимог, щоб оцінити, наскільки якість набору даних відповідає його очікуванням. Піделементи якості даних включають горизонтальну та вертикальну точність для позиційної точності; точність класифікації (об’єктивної та атрибутивної) і кількісної точності атрибутів для тематичної точності; часову точність, часову узгодженість і часову дійсність для часової характеристики; повноту об’єктів, атрибутів і просторову повноту для категорії повноти; а також узгодженість доменів, форматів і топологічну узгодженість для логічної узгодженості.

Виробник даних визначає обсяг оцінки якості для кожного піделемента, необхідного для перевірки відповідності специфікації та виконання поставленої мети продукту. Обсяг може відрізнятися залежно від специфікації продукту, і кожен піделемент якості даних може мати різни обсяг. Відштовхуючись від вимог користувача, користувач даних встановлює необхідний обсяг оцінки якості, щоб перевірити, наскільки якість набору даних відповідає цим вимогам. Для кожного відповідного піделемента якості даних надається один обсяг якості. Обсяг якості може відноситися до серії наборів даних, до якої належить набір, самого набору даних або визначеної групи звітності. Специфікація продукту та оглядові

елементи якості даних використовуються для визначення обсягу якості даних для кожного відповідного піделемента.

Якість може варіюватися в межах одного набору даних. Для більш повного відображення інформації про якість можуть бути надані кілька обсягів якості для кожного відповідного піделемента. Кожен обсяг якості повинен бути чітко визначений. Для цього можна використовувати наступне:

- рівень, наприклад, серія наборів даних, сам набір даних або група звітності;
- типи елементів (перелік типів об’єктів та їхніх атрибутів) або конкретні елементи (перелік випадків об’єктів та значень атрибутів);
- географічний охоплення;
- часовий діапазон, зокрема часові рамки та їхня точність.

Вибір міри оцінки якості здійснюється відповідно до обраного обсягу. Для кожного обсягу якості даних надається одна міра оцінки якості. Водночас одного показника може бути недостатньо для повної оцінки якості та забезпечення відповідності набору даних усім можливим сценаріям використання. Комбінація кількох мір оцінки якості може дати більш інформативні результати. Для кожної обраної міри оцінки якості виробник або користувач обирає метод оцінювання якості. Існує два методи: прямий і непрямий. Прямий метод виконується шляхом вибіркового аналізу або повної перевірки даних у наборі, тоді як непрямий метод — через оцінку інформації про якість з інших джерел, окрім самого набору. Вибір методу оцінювання визначається як виробником, так і користувачем даних.

Існують два типи методів прямої оцінки: метод повної перевірки та метод вибіркової оцінки. Метод повної перевірки передбачає тестування 100% об'єктів однієї сукупності для визначення якості. Метод вибіркової оцінки включає перевірку лише вибірки із загальної кількості об’єктів для оцінки їхньої якості. Як вибірковий метод, так і метод повної перевірки можуть бути реалізовані автоматизованими або неавтоматизованими засобами. Наприклад, деякі типи помилок, такі як узгодженість ознак, узгодженість атрибутів або значень

атрибутів, можуть виявлятися автоматично за допомогою комп’ютерних програм.

Для вибіркової перевірки можуть бути використані стандарти ISO 2859 (Перевірка за атрибутами) та ISO 3951 (Перевірка за змінними) [11,12]. Основні етапи вибіркової перевірки включають:

1) Визначення об’єктів перевірки. Об’єктом може бути будь-який компонент, який можна індивідуально описати або оцінити, наприклад, певна характеристика, атрибут характеристик або їх комбінація.

2) Розподіл обсягу даних за статистично однорідними партіями. Однорідність може оцінюватися за джерелом даних виробництва, системою виробництва (апаратне забезпечення, програмне забезпечення, кваліфікація оператора), складністю та густотою характеристик. Партія – це мінімальна одиниця, до якої прив'язаний результат оцінки якості. Якщо партія не проходить перевірку, усі об’єкти з цієї партії можуть бути відхилені або перероблені.

3) Розбиття партій на вибіркові одиниці. Вибіркова одиниця є найменшою географічною зоною у моделі світу, для якої виконується перевірка всіх об’єктів, що належать до неї. Кількість об'єктів у партії вважається її розміром. За допомогою простого випадкового відбору для перевірки здійснюють вибір вибіркових одиниць із кожної партії.

4) Перевіряються всі об’єкти, що входять до вибраних вибіркових одиниць.

5) Якщо кількість помилкових об’єктів досягає певного встановленого порогу відхилення, партія не приймається. Оскільки перевірка проводиться за атрибутами, застосовується стандарт ISO 2859 [11]. Якщо середнє значення перевірених значень не відповідають граничним умовам, партія також не приймається.

6) Якщо всі партії проходять перевірку, то обсяг даних вважається прийнятним.

Непрямі методи оцінювання базуються на оцінках показників якості даних, отриманих із джерел, відмінних від самих елементів даних у наборі. Різноманіття

таких джерел включає, але не обмежується, метаданими, знаннями про призначення набору даних, документацією про його походження, історією використання набору даних та звітами щодо якості даних, що застосовувались для створення набору. Також корисною є інформація про процес створення та можливі помилки, які були допущені або виявлені у ході цього процесу.

У межах даного дослідження практична реалізація методики оцінювання якості відкритих геопросторових даних виконується шляхом порівняння інформації з п’ятьох джерел: OpenStreetMap (OSM), Microsoft Bing Building Footprints, Google Satellite, Bing Maps і Esri Imagery.

Еталоном для порівняння виступають топографічні плани масштабу 1:2000. Усі набори даних переведено в єдину систему координат УСК-2000 (EPSG:5561). Для кожного населеного пункту створюються дві групи шарів для аналізу:

1) Еталонні дані:
   - Полігональні об'єкти (будівлі, представлені на топографічному плані).
   - Точкові об'єкти (вершини контурів будівель).
2) Оцінювані набори даних:
   - Полігони, отримані з кожного джерела.
   - Відповідні точки контурів будівель.

Для оцінювання якості відкритих джерел застосовуються такі ключові показники:

1) Позиційна точність – розраховується на основі координат вершин полігональних об'єктів через визначення середньоквадратичної похибки (СКП) між точками еталону і оцінюваних джерел.

2) Повнота – показує, який відсоток об'єктів з еталонного шару представлений у кожному відкритому джерелі.

3) Детальність – оцінює, наскільки площа або конфігурація будівель (полігонів) у джерелах збігається з еталонними даними.

Для аналізу кожного набору відкритих геопросторових даних формується набір із 30 полігональних об’єктів будівель та споруд, які містяться в усіх наборах

одночасно. Оцінка здійснена за допомогою методів геопросторового та статистичного аналізу, а результати подані у відповідних таблицях.

Також виконано аналіз розподілу похибок для визначення їх природи – випадкової або систематичної. Під час систематичних відхилень використовується: формула Гаусса (2.1) – коли істинне значення відоме, та формула Бесселя (2.2) – коли істинне значення невідоме:

$$m = \sqrt{\frac{1}{n} \cdot \sum_{i=1}^{n} \Delta_i^2} \tag{2.1}$$

де i = 1…n;

n – кількість вимірів;

Δ_i – похибка і-го вимірювання, яка визначається як різниця між виміряним значенням координати та еталонним (істинним) значенням координати.

$$\Delta_i = x_i^{\text{оцін}} - x_i^{\text{еталон}}$$

Ця формула використовується в тому випадку, якщо відома точна координата контрольної точки (еталонне значення), і кожне виміряне значення порівнюється з нею. Формула дозволяє визначити середньо квадратичну похибку між виміряними і справжніми координатами, що є показником абсолютної точності геопросторових даних.

$$m = \sqrt{\frac{1}{n-1} \cdot \sum_{i=1}^{n} \delta_i^2} \tag{2.2}$$

де i = 1…n;

δ_i – різниця між виміряним значенням та середнім арифметичним значенням даної величини.

$$\delta_i = x_i - \bar{x}$$

де x_i – виміряне значення координати;

$\bar{x}$ – середнє арифметичне всіх виміряних значень.

Ця формула використовується, коли істинне значення невідоме, і для оцінки похибок застосовується середнє значення як умовне "еталонне". Вона дозволяє врахувати внутрішню варіативність вимірювань та уникнути зміщення в оцінці дисперсії.

Під час реалізації оцінювання якості відкритих геопросторових даних здійснено низку послідовних етапів, які забезпечили комплексний підхід до аналізу:

1) Визначення джерел та об’єктів дослідження. Були обрані п’ять популярних джерел відкритих даних – OpenStreetMap, Bing Maps, Google та Esri. Для аналізу відібрано ті самі об'єкти (будівлі), наявні у всіх джерелах та на топографічному плані масштабу 1:2000.

2) Підготовка еталонних моделей геопросторових об’єктів. Контрольні дані були отримані шляхом експорту полігональних об’єктів будівель та споруд (контурів) із цифрового топографічного плану масштабу 1:2000, на основі чого були створені додатково точкові моделі (вершини контурів).

3) Формування оцінюваних моделей геопросторових об’єктів. Для кожного набору створено відповідні шари у ГІС-проєкті: полігональний (будівлі) і точковий (вершини). Дані приведено до єдиної системи координат для коректного порівняльного аналізу.

4) Уніфікація вибірки. Загальна вибірка для аналізу сформована з 30 будівель, які є в усіх наборах. Це дозволило забезпечити рівні умови оцінювання.

5) Проведення просторового аналізу. Здійснено порівняння координат точок для обчислення позиційної точності, а також аналіз відповідності площ та геометрії контурів визначеним вимогам щодо детальності та повноти даних відповідним масштабам топографічних планів і карт.

6) Обчислення кількісних показників. Для кожного об’єкта визначено значення трьох показників якості: точність (за СКП планового положення вершин полігональних об’єктів), повнота (відповідність мінімальним розмірам полігональних об’єктів), детальність (відповідність максимальній відстані між точками векторних моделей контурів об’єктів залежно від масштабу

топографічної зйомки, що визначає роздільну здатність векторних моделей контурів об’єктів).

7) Статистичне опрацювання результатів. Виконано обчислення розподілу різниць між еталонними та виміряними значеннями, визначено СКП планового положення полігональних об’єктів для кожного набору. Визначено характер відхилень: випадкові або систематичні.

8) Інтерпретація результатів. Проведено аналіз придатності кожного набору до застосування в межах конкретного масштабу топографічних планів і карт. Сформовано висновки щодо доцільності використання досліджених наборів відкритих геопросторових даних у різних сферах.

2.2 Концептуальна модель бази геопросторових даних для оцінювання якості даних

Ефективне оцінювання якості геопросторових даних потребує наявності структурованої інформаційної системи, яка включає не лише самі дані про об’єкти, але й метадані, показники якості, результати перевірок, часові характеристики та джерела походження. Для цього доцільно розробити концептуальну модель бази геопросторових даних, яка забезпечить цілісність, повноту та стандартизоване подання інформації про якість.

Відповідно до стандарту ISO 19157, передбачається створення в рамках технічних вимог вхідної моделі якості НГД (рис. 2.1) [6]. Ця модель надалі застосовується переважно для оцінювання внутрішньої якості продукції.

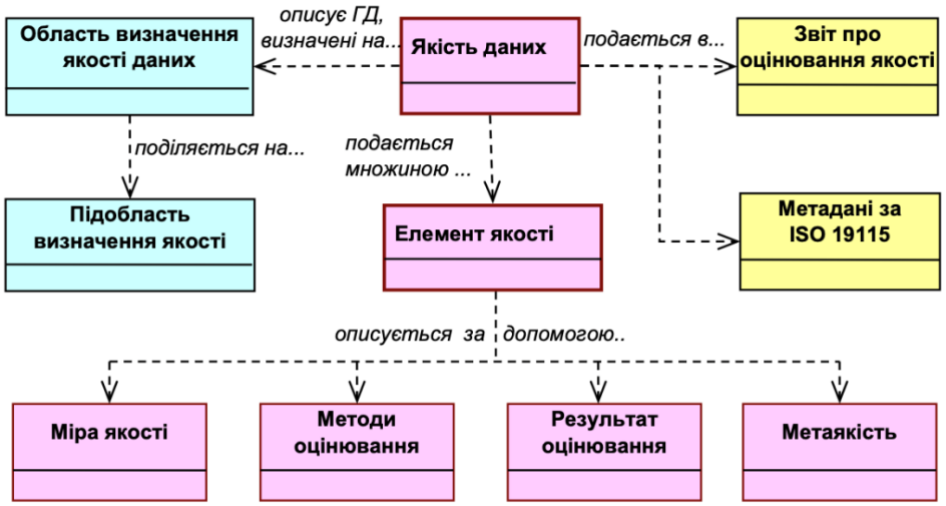


Рис 2.1 – Концептуальна модель якості геопросторових даних за ISO 19157[6]

Згідно з концептуальною моделлю, якість даних представлена як сукупність взаємозалежних елементів. Кожен із цих елементів характеризується кількома ключовими аспектами, серед яких міра якості, методи оцінювання, отримані результати оцінки, а також метаякість, що є метаданими про якість самих даних. Такий підхід дозволяє систематично аналізувати якість інформації. Характеристика якості стосується певної сукупності геопросторових даних. У структурі концептуальної моделі це відображено у вигляді області визначення якості, яка може бути деталізована на менші підобласті. Це означає, що як вимоги до якості, так і методологія її оцінки повинні розглядатися не тільки загалом для всього набору даних, але й окремо для його складових частин. Складові можуть включати різні типи об’єктів, специфічні тематичні та просторові атрибути, асоціації або обмеження, що визначені у специфікації відповідного геопросторового набору. Такий підхід виправданий тим, що в межах одного набору даних можуть існувати різні вимоги до просторової точності положення об’єктів різних типів або розрізнена точність атрибутів залежно від їхньої вагомості у вирішенні конкретних прикладних завдань. На основі концепції оцінювання якості даних визначено, що будь-яка сутність набору даних, зокрема екземпляри типів об'єктів, значення атрибутів або види взаємозв’язків між компонентами, можуть мати свої окремі характеристики якості. Крім того, якість

окремих частин набору даних може суттєво відрізнятися від загальної якості всього набору. Для цього концепція передбачає можливість отримувати інформацію не лише про якість набору в цілому, але й про якість його дрібніших структурних складових. Це відображає особливості геопросторових даних та уможливлює більш точний контроль і аналіз. Для спрощення доступу до інформації про якість даних ця інформація включається до спеціальних звітів та відповідних метаданих. Звіт про якість є обов’язковим підсумковим документом за результатами процесу оцінювання. При цьому використовуються специфіковані процедури та методи, що закладені у базову модель якості. Зазвичай такі звіти служать основою для подальшого виявлення і корекції помилок у наборах даних, знайдених під час контролю якості. Окремо інформація щодо якості інтегрується у відповідні розділи метаданих. Ці метадані стають невід'ємною частиною кінцевого дата-продукту та можуть використовуватися споживачами у процесі аналізу зовнішньої якості отриманих даних. Таким чином, уся концепція призначена для забезпечення комплексного підходу до управління якістю даних шляхом деталізації кожного елемента і максимального врахування специфічних вимог до окремих компонентів геопросторового набору. Це дає змогу розробникам і користувачам досягати високого рівня точності та надійності кінцевого продукту.

Для систематизації та уніфікації просторових об’єктів, що підлягають оцінюванню, розроблено концептуальну модель бази геопросторових даних (рис. 2.2). Модель побудована з використанням сутностей, які відображають основні класи об’єктів, їхні атрибути та взаємозв’язки, що забезпечують можливість зберігання, порівняння та аналізу якості просторових даних.

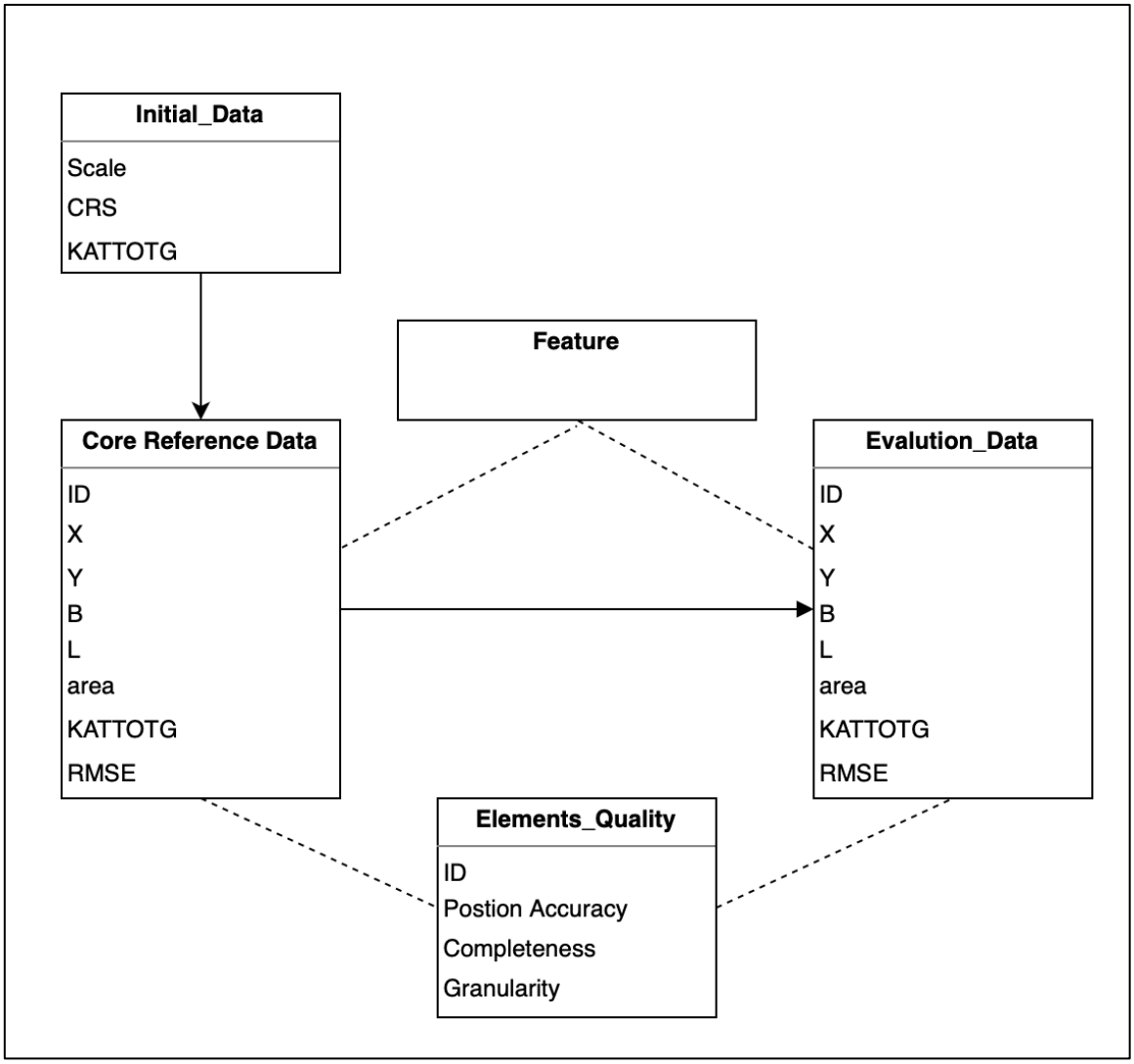


Рис. 2.2 – Концептуальна модель бази геопросторових даних

У центрі моделі розміщено абстрактний клас «Feature», який слугує узагальненою структурою для двох основних типів об’єктів: еталонних (вихідних з топографічного плану) та оцінюваних (отриманих із відкритих джерел). Обидва ці класи наслідують основні характеристики, визначені у базовому класі, що дозволяє застосовувати єдиний підхід до оцінки їх якості.

Кожен об’єкт містить набір просторових (координати X, Y, параметри B, L, площа) та описових (ідентифікатор ID, код території – KATTOTG, середньоквадратична похибка – RMSE) характеристик. Така структура дозволяє точно визначати розташування та властивості об’єктів у межах досліджуваної території.

Зв’язок між еталонним шаром та джерелом його формування реалізовано через клас «Initial_Data», що містить параметри масштабу, системи координат і посилання на KATTOTG. Це дозволяє враховувати контекст створення базових даних при порівнянні з іншими джерелами.

Окремим елементом моделі виступає клас «Elements_Quality», який пов’язаний з обома типами об’єктів. У ньому фіксуються три ключові критерії: позиційна точність, повнота та детальність. Саме ці показники використовуються для формалізованої оцінки відповідності відкритих геопросторових даних до еталонної інформації.

Таким чином, концептуальна модель дозволяє чітко структурувати просторові об’єкти, забезпечує основу для їх порівняння, а також формує засади для подальшого впровадження методів контролю якості геопросторових даних на етапах збору, обробки та використання.

2.3 Логічна модель бази геопросторових даних для оцінювання якості даних

Логічна модель – це схема організації інформації, яка визначає, які саме дані повинні зберігатися в базі, як вони взаємопов’язані та які атрибути кожна сутність повинна містити. Вона є проміжною ланкою між загальною ідеєю (концептуальною моделлю) та її реалізацією в конкретній системі. Така модель дозволяє чітко описати склад і структуру просторових об’єктів, а також встановити правила їхнього подання у вигляді таблиць, полів і зв’язків.

Логічна модель бази геопросторових даних визначає структуру зберігання просторових об'єктів та атрибутивної інформації, необхідної для оцінювання якості відкритих геопросторових джерел (рис. 2.3). Вона створена на основі концептуальної моделі та адаптована для аналізу точності, повноти і детальності.

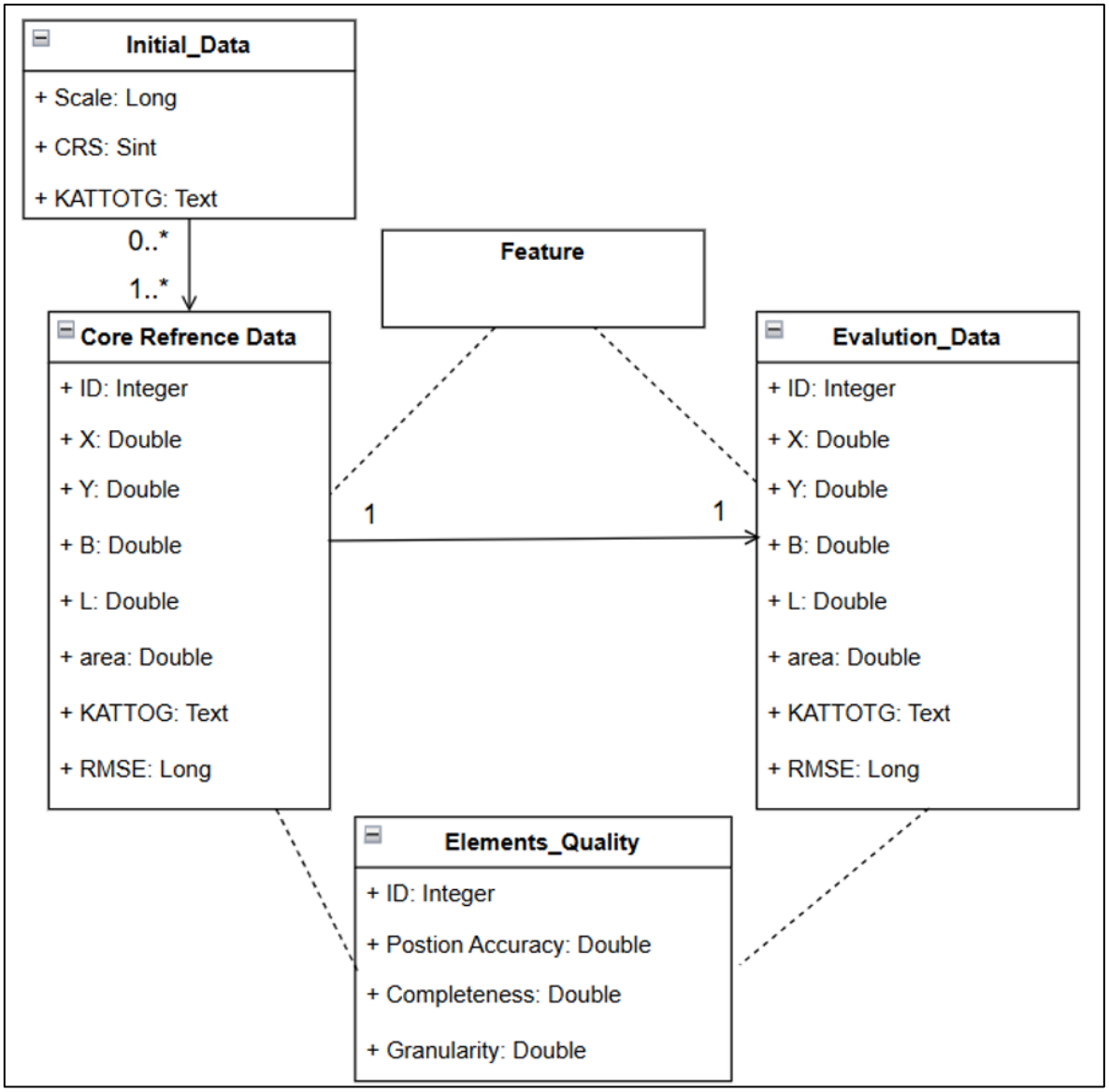


Рис. 2.3 – Логічна модель бази геопросторових даних

На відміну від концептуального рівня, де моделюються лише загальні об’єкти та їхні зв’язки, логічна модель уточнює, які поля має містити кожна таблиця, які типи даних використовуються та як забезпечити узгодженість між еталонними і оцінюваними об’єктами. Завдяки цьому створюється основа для зручного зберігання, фільтрації, порівняння та аналізу просторових даних у середовищах типу ArcGIS.

Кожен просторовий об’єкт у базі має ряд параметрів, які забезпечують його ідентифікацію, локалізацію в просторі та опис геометричних характеристик. Ідентифікатор ID використовується для унікального розпізнавання об’єкта і

зв'язку з іншими таблицями. Координати X та Y зберігаються у прямокутній системі координат УСК-2000, слугуючи основою для просторових обчислень. Географічні координати B (широта) і L (довгота), представлені у системі WGS 84, застосовуються для трансформацій.

Атрибут area містить значення площі полігонального об'єкта, розраховане автоматично в квадратних метрах для оцінки відповідності геометрії еталонним даним. Атрибут KATTOTG відображає код територіальної громади, до якої належить об'єкт, що забезпечує територіальну прив'язку і групування даних. Показник RMSE (середньоквадратична похибка) обчислюється під час аналізу координат об'єкта стосовно еталонного шару та визначає його позиційну точність.

Окрім основної таблиці об'єктів, модель включає додаткову таблицю результатів оцінки якості. У ній зберігаються показники точності, повноти та детальності для кожної пари даних: відкритого джерела та відповідного еталону. Ці об'єкти пов'язані через ID, а оцінювання здійснюється автоматичним порівнянням координат, площ, кількості вершин чи інших геометричних характеристик.

Також у моделі передбачена таблиця з даними про джерела еталонної інформації – топографічний план. Вона містить відомості про масштаб, систему координат і прив'язується до об'єктів за атрибутом KATTOTG.

Розроблена логічна модель підходить для реалізації у середовищі ArcGIS. Вона забезпечує просторову цілісність даних, дозволяє виконувати точні розрахунки якості, створювати звіти та здійснювати аналіз результатів для різноманітних джерел геопросторової інформації.

2.4 Опрацювання вихідних геопросторових даних проєкту

Для проведення оцінки якості відкритих геопросторових даних було виконано комплексне опрацювання вхідної інформації з різних джерел. Вся підготовка проводилася у середовищі ArcGIS Desktop, з використанням

можливостей програм ArcMap та ArcCatalog. Основною метою цього етапу було узгодження даних, приведення їх до єдиної системи координат і формування структур, необхідних для подальшого аналізу точності.

До складу використаних джерел увійшли:

1) векторні шари з OpenStreetMap (OSM);
2) дані з Microsoft Bing Building Footprints;
3) супутникові зображення із сервісів Google, Bing та Esri;
4) топографічні плани масштабу 1:2000 на населені пункти: Тульчин, Славута, Лисянка, Пісочин та Рогань, які були обрані як еталонні шари.

Усі дані було приведено до УСК-2000 (EPSG:5561) – системи координат, яка є чинною в Україні. Зміна проекції здійснювалась за допомогою інструменту Project, який дозволяє перетворити координати геопросторових об’єктів з допустимою точністю.

Нові шари було створено в ArcCatalog. Під час створення шейп-файлів задавались необхідний тип геометрії (точки або полігони) та вказувалась проекція. Подальше редагування здійснювалось у ArcMap, де об’єкти створювались вручну або за допомогою оцифрування на основі підкладки (наприклад, топоплану чи супутникових знімків).

У результаті було сформовано:

1) контрольні полігони будівель – створені вручну на основі еталонного топографічного плану;
2) контрольні точки – вершини цих полігонів;
3) оцінювані полігони – будівлі, отримані з наборів відкритих геопросторових даних;
4) оцінювані точки – геометричні вершини оцінюваних об’єктів.

Перед початком аналізу всі об'єкти були перевірені на наявність геометричних помилок. Застосовувались стандартні інструменти ArcGIS для виявлення некоректних елементів – самоперетинів, незамкнених контурів або пустих геометрій. Для корекції таких помилок використовувалися Check Geometry та Repair Geometry.

Обробка охоплювала території кількох населених пунктів – Тульчина (адміністративний центр Тульчинського району Вінницької області та Тульчинської міської громади), Славути (місто в Шепетівському районі Хмельницької області), Рогані (адміністративний центр Роганської селищної громади Харківського району Харківської області), Пісочина (селище Харківського району Харківської області) та Лисянки (селище міського типу Черкаської області). Для відсікання зайвих даних використовувався інструмент Clip, який дозволяє вирізати просторові об’єкти в межах заданої області.

Щоб мати змогу аналізувати координати, у таблиці атрибутів було створено поля X та Y. Значення координат обчислювались через функцію Calculate Geometry, при цьому координати зберігалися у метрах у межах УСК-2000. У разі потреби координати центроїдів або вершин визначалися за допомогою функцій перетворення геометрії.

Підготовлені дані експортувалися до Excel (у форматі .dbf або .csv), де на основі координат контрольних і оцінюваних точок розраховувалися похибки за формулами Гаусса і Бесселя. Ці значення використовувались для визначення просторової точності та подальшого аналізу результатів оцінки якості.

Таким чином, на цьому етапі було створено єдину узгоджену базу геоданих, що стала основою для аналітичного оцінювання точності, повноти та геометричної відповідності відкритих джерел.

Висновок до розділу 2: У цьому розділі розроблено методику, яка комплексно оцінює якість відкритих геопросторових даних на основі вимог стандарту ISO 19157. Пропонується аналізувати дані за трьома основними показниками: позиційною точністю, повнотою та детальністю. Створено концептуальну та логічну моделі бази геоданих, що забезпечують зручне зберігання, доступ і обробку інформації для оцінки якості. Також проведено підготовку вихідних геопросторових джерел із врахуванням їх структури, формату, систем координат і походження, що дозволяє здійснити подальший порівняльний аналіз і забезпечити об'єктивність оцінювання.

РОЗДІЛ 3. ДОСЛІДНА РЕАЛІЗАЦІЯ МЕТОДИКИ ОЦІНКИ ЯКОСТІ ВІДКРИТИХ ГЕОПРОСТОРОВИХ ДАНИХ

3.1 Реалізація бази геопросторових даних проєкту у середовищі ArcGIS

На цьому етапі було створено базу геопросторових даних, яка є основою для зберігання, організації та опрацювання просторової інформації проєкту. Формат File Geodatabase (GDB) було обрано завдяки його здатності стабільно працювати з великими обсягами даних, підтримувати складні типи об’єктів та забезпечувати зручне управління атрибутивною структурою.

Базу геоданих створено в середовищі ArcCatalog, яке є частиною платформи ArcGIS Desktop. Через меню New → File Geodatabase було створено нову базу з назвою «Qualityssessment.gdb», що відповідала тематиці дослідження. Усередині бази для кожного досліджуваного населеного пункту, таких як Тульчин, Славута, Лисянка, Пісочин і Рогань, сформовано окремі Feature Dataset. Усі вони реалізовані в проєкції УСК-2000 (EPSG:5561), яка є офіційною системою координат України, що уніфікує просторову взаємодію об’єктів. Дата цієї системи координат була додана вручну, оскільки актуальні параметри відсутні в ArcGIS for Desktop.

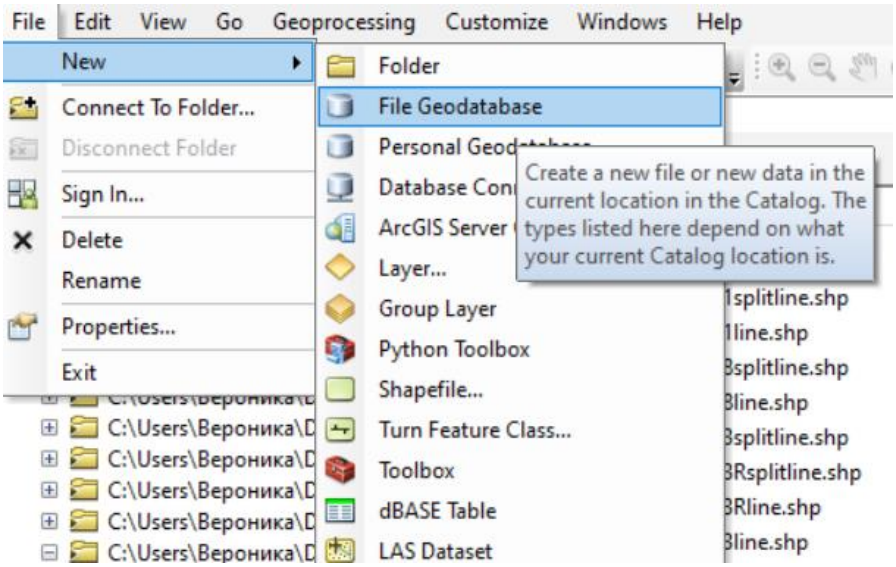


Рис. 3.1 – Створення бази геоданих (GDB)

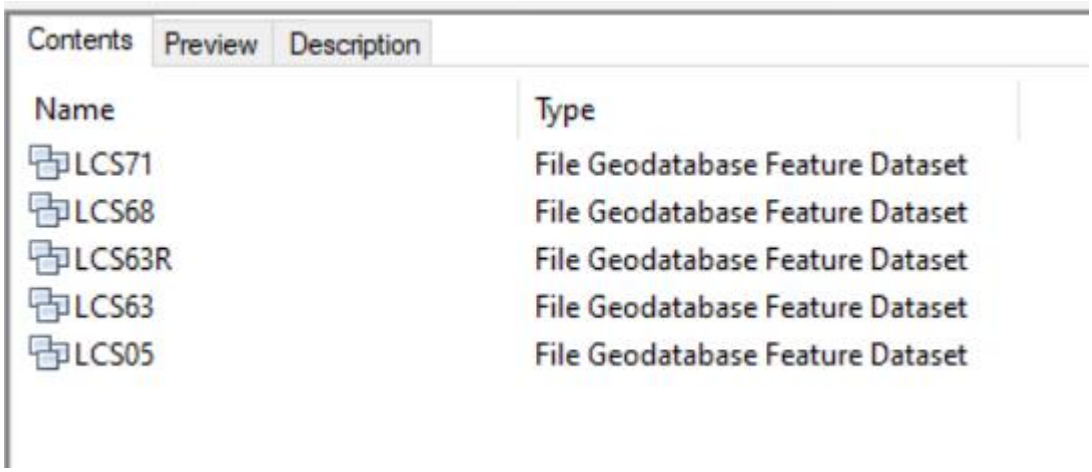


Рис. 3.2 – Створення Feature Dataset

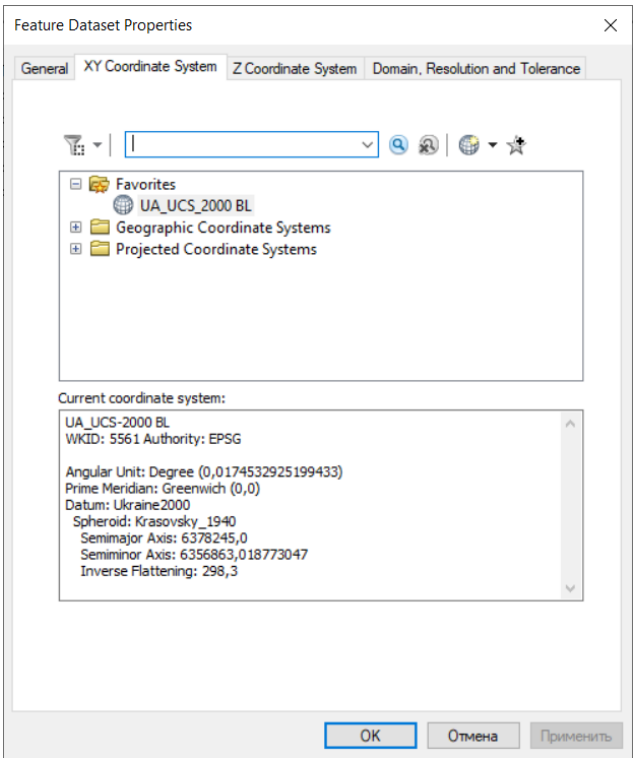


Рис. 3.3 – Додавання системи координат

Кожен Feature Dataset містить класи об’єктів (Feature Classes), що виконують функцію контейнерів просторових даних. Для кожного класу визначений відповідний тип геометрії: полігони використовувалися для зберігання контурів будівель, точки – для їхніх вершин. Під час створення кожного шару вручну налаштовувалася структура атрибутивної таблиці відповідно до логічної моделі даних. Основні поля включали ID – унікальний ідентифікатор об’єкта; X і Y – метричні координати; B і L – географічні

координати за широтою та довготою; area – площу об’єкта; ktotg – код територіальної громади; СКП – середньоквадратичну похибку.

Після завершення налаштування структура бази геоданих готова до наповнення об’єктами та виконання подальшого дослідження. Однією з найважливіших переваг середовища ArcGIS є централізоване управління всіма проєктними шарами, структурована організація інформації за територіальним принципом та забезпечення просторової цілісності між об’єктами. Ця GDB є платформою для подальшого аналізу якості відкритих геопросторових джерел через інтеграцію еталонних і оцінюваних даних у єдину систему.

3.2 Підготовка ГІС-проекту для оцінювання якості відкритих геопросторових даних

Після створення структури бази геоданих наступним важливим етапом стало наповнення її реальними просторовими об'єктами, зокрема будівлями, які були отримані з еталонних джерел (топографічний план) та відкритих платформ (OpenStreetMap, Microsoft Building Footprints, Bing, Google, Esri). Для виконання цього завдання використовували два основні методи: імпорт заздалегідь підготовлених шарів і ручне оцифрування нових об’єктів.

Імпорт здійснювали за допомогою інструментів Feature Class to Feature Class, які забезпечували додавання просторових даних до існуючих класів без порушення їхньої структури. Всі шари, отримані з відкритих джерел, заздалегідь конвертували у формат .shp або .csv, перевіряли на відповідність системі координат УСК-2000 і налаштовували так, щоб вони мали мінімально вимогливу структуру атрибутів.

Для даних, які створювали вручну, використовували інструменти редагування в середовищі ArcMap. Оцифровування кожної будівлі здійснювали на основі підкладки (топографічного плану або супутникового знімку), після чого вручну вписували значення в таблицю атрибутів. Координати та площі об’єктів визначалися автоматично через функцію Calculate Geometry. Показники B і L

отримували шляхом трансформації координат у географічну систему WGS 84, а поле СКП заповнювали на основі аналізу похибок, який проводився поза межами ГІС-системи.

Кожен Feature Class у проєкті містив повний набір об’єктів із відповідними атрибутами, що дозволяло проводити просторове порівняння, перевіряти відповідність геометрії, оцінювати кількість збігів із еталоном і розраховувати точні показники якості даних. Використання єдиної геобази забезпечувало зручний доступ до даних, можливість формування запитів, візуального аналізу та створення картографічних результатів.

Таким чином, у рамках підготовки ГІС-проєкту не тільки наповнювали базу геоданих просторовими об’єктами, але й створювали інтегроване інформаційне середовище для автоматизованого аналізу відкритих геопросторових джерел за формалізованими критеріями.

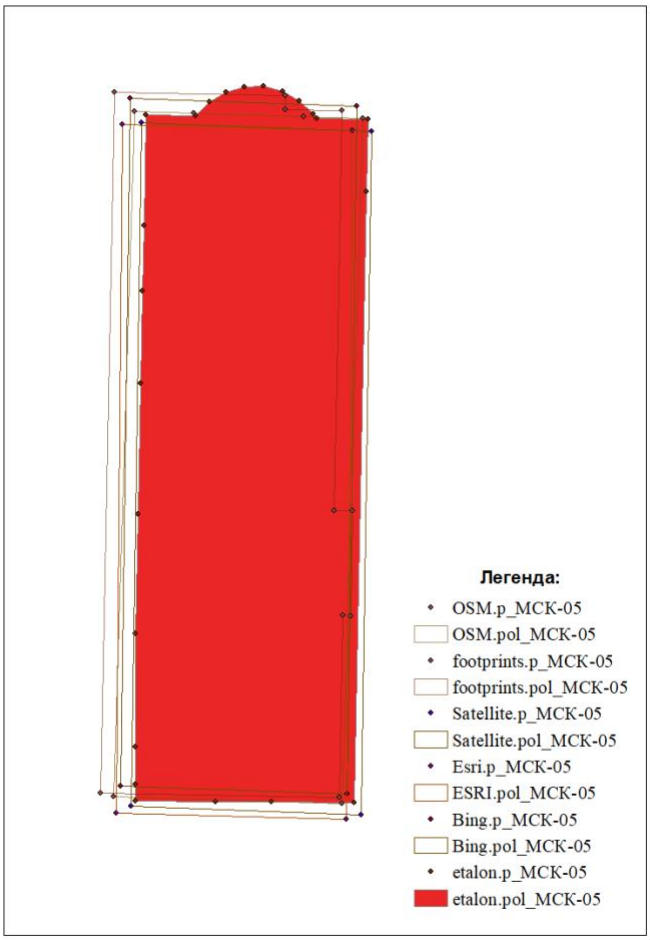


Рис. 3.4 – Порівняльна візуалізація геометрії будівель з різних джерел

На ілюстрації (рис. 3.4) зображено приклад одного будівельного об’єкта, межі якого були отримані з кількох відкритих джерел геопросторової інформації. Для забезпечення коректного просторового зіставлення всі шари приведено до єдиної локальної системи координат МСК-05.

У порівнянні задіяно такі шари:

– OSM.p_МСК-05 / OSM.pol_МСК-05 – точкові та полігональні дані з OpenStreetMap;

– footprints.p_МСК-05 / footprints.pol_МСК-05 – будівлі з набору Microsoft Building Footprints;

– Satellite.p_МСК-05 / Satellite.pol_МСК-05 – контури, векторизовані зі знімків Google Satellite;

– Esri.p_МСК-05 / ESRI.pol_МСК-05 – геометрія будівель зображена на супутникових знімках Esri;

– Bing.p_МСК-05 / Bing.pol_МСК-05 – контури, отримані з Bing Maps;

– etalon.p_МСК-05 / etalon.pol_МСК-05 – еталонні дані, побудовані на основі топоплану масштабу 1:2000.

Червоним кольором позначено еталонний полігон, який слугує базовим для оцінювання просторової якості. Інші контури подано у вигляді напівпрозорих шарів з відображенням вершин, що дозволяє візуально оцінити:

1. Позиційну точність (точність розташування)
2. Повноту (відсутність або наявність окремих частин)
3. Детальність (форма, кількість і довжина сегментів)

Такий спосіб представлення даних дозволяє візуально виявити відмінності між джерелами та є підґрунтям для подальшого кількісного аналізу просторової відповідності.

3.3 Виконання оцінки якості відкритих геопросторових даних у середовищі ArcGIS

У процесі виконання цього підрозділу здійснено оцінювання просторової якості геоданих із різних відкритих джерел у середовищі ArcGIS Desktop. Аналіз проводився за трьома ключовими показниками: позиційна точність, повнота та детальність геометрії. Для реалізації завдання було створено необхідну базу геоданих, підготовлено ГІС-проєкт і виконано низку аналітичних процедур.

Позиційна точність:

На цьому етапі визначити позиційну точність не вдалося, оскільки об’єкти, отримані з відкритих джерел геопросторових даних, мають суттєві відмінності у порівнянні з еталонними. Зокрема, контури будівель різняться за кількістю вершин та конфігурацією, що унеможливлює точне порівняння координат відповідних точок. Через це автоматичне зіставлення точкових елементів виявилося недоцільним, адже результати були б некоректними або неповними.

У зв’язку з цим, у рамках даної роботи позиційна точність не оцінювалася, а подальший аналіз зосереджувався на повноті та детальності геометрії будівель.

Детальність:

У межах дослідження було проведено оцінку детальності геометрії будівель, представлених у різних наборах відкритих геопросторових даних. Основною метою було визначення, наскільки форма будівель зберігає геометричну точність і відповідність детальності, необхідній для використання в масштабі 1:2000. Як базовий критерій застосовано граничне значення довжини сторони – 40 метрів, відповідно до нормативів роздільної здатності векторних моделей.

Для аналізу використовувалась наступна послідовність дій:

1. Перетворення полігонів у лінійні об’єкти. Полігональні шари будівель були оброблені за допомогою інструменту «Feature to Line» у середовищі ArcGIS. Це дозволило перейти від суцільних контурів до окремих лінійних елементів – сторін будівель.

2. Розбиття ліній на сегменти. Застосовано інструмент «Split Line at Vertices», який розділив контури на відрізки між вершинами. Таким чином, кожна сторона будівлі була представленою як окремий сегмент, придатний для індивідуального аналізу.

3. Обчислення довжин сегментів. Для кожного лінійного сегмента була розрахована довжина у метрах за допомогою функції Calculate Geometry.

4. Класифікація об’єктів за рівнем деталізації. На основі аналізу довжин усіх сторін кожної будівлі було здійснено розподіл об’єктів на три категорії:

– Повністю відповідає – усі сторони будівлі не перевищують 40 м;

– Частково відповідає – деякі сторони перевищують 40 м, але не всі;

– Не відповідає повністю – всі сторони мають довжину понад 40 м.

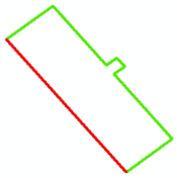
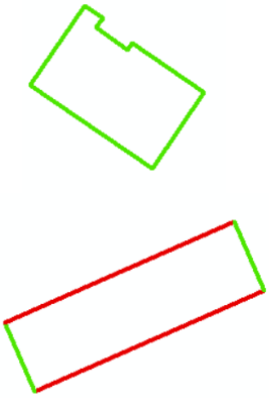
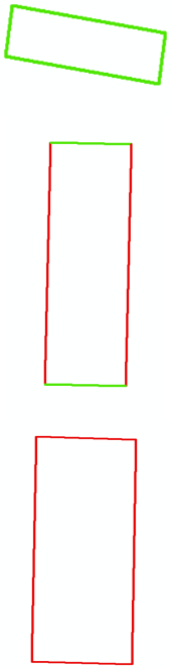
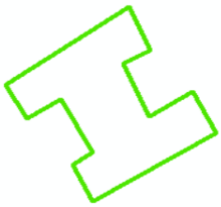
5. Візуалізація результатів. Для покращення сприйняття та виявлення об’єктів з недостатньою деталізацією, сегменти було зафарбовано кольорами: зелений – сторони, що відповідають вимогам; червоний – ті, що їх перевищують. Це дозволило швидко і наочно ідентифікувати проблемні об’єкти.

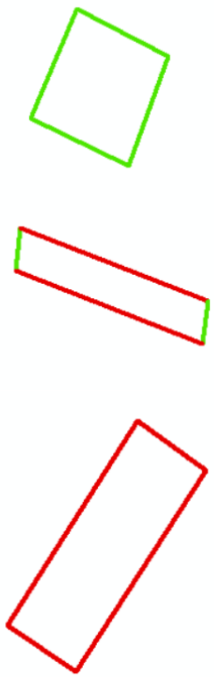
6. Формування зведеної таблиці. Для кожного джерела відкритих геоданих було створено узагальнену таблицю, в якій зазначено кількість об’єктів, що повністю, частково або зовсім не відповідають критерію детальності. Також для кожної категорії додано візуальні приклади для кращого розуміння результатів.

Таблиця 3.1

Оцінка детальності геометрії будівель за наборами відкритих геопросторових даних

Набір відкритих даних	Повністю відповідає (к-сть)	Частково відповідає (к-сть)	Не відповідає повністю (к-сть)	Візуальні приклади об’єктів
OSM	31	92	0	

				
Microsoft Building Footprints	83	65	0	
Google Satellite	88	59	2	
Esri Imagery	84	63	2	

Bing Maps	86	61	2	

Загалом було проаналізовано близько 150 об’єктів з кожного набору відкритих геопросторових даних. В окремих випадках кількість була меншою (наприклад, 149 або 148), що пов’язано зі складністю або неможливістю однозначної векторизації деяких об’єктів у певних джерелах.

За результатами відсоткового аналізу встановлено:

– Google Satellite – 59,1% об’єктів повністю відповідають вимогам детальності, 39,6% – частково відповідають, 1,3% – повністю не відповідають.

– Bing Maps – 57,7% повністю відповідають вимогам детальності, 40,9% – частково відповідають, 1,3% – не відповідають.

– Esri Imagery – 56,4% повністю відповідають вимогам детальності, 42,3% – частково відповідають, 1,3% – не відповідають.

– Microsoft Building Footprints – 56,1% об’єктів повністю відповідають вимогам детальності, 43,9% – частково відповідають, повністю непридатних об’єктів не виявлено.

– OpenStreetMap (OSM) – лише 25,2% повністю відповідають критерію, решта 74,8% – лише частково відповідають, повністю непридатних об’єктів не зафіксовано.

Отже, найбільш деталізованими та збалансованими щодо геометрії виявилися Google Satellite, Bing, Esri та Microsoft, тоді як OSM продемонстрував нижчий рівень деталізації, що обмежує його придатність для високоточних завдань.

Повнота:

Під повнотою у контексті відкритих геопросторових даних розуміють здатність джерела передавати максимально повну інформацію про наявні об’єкти на території дослідження. У даному випадку повнота оцінювалась шляхом порівняння площ будівель, отриманих з різних джерел, з референтними значеннями, а також аналізу розмаху (мінімальної та максимальної площ) об’єктів, що дозволяє встановити орієнтовний масштаб використання даних.

Таблиця 3.2

Аналіз повноти геопросторових даних за мінімальною та максимальною площею будівель

Набір даних	Мін. площа ($м^2$)	Макс. площа ($м^2$)	Орієнтовний масштаб

Microsoft Building Footprints	0,2542	28983,8289	1:500 -1:50000
OSM	7,5542	28840,9448	1:2000 -1:50000
Google Satellite	59,8087	9717,2639	1:5000 -1:50000
Esri Imagery	59,1054	9763,0554	1:5000 -1:50000
Bing Maps	45,2426	9624,4272	1:5000- 1:50000

Для кожного набору даних визначались:

1. Мінімальна площа об’єкта – відображає здатність джерела фіксувати дрібні, незначні за площею будівлі.

2. Максимальна площа об’єкта – показує, чи були належним чином охоплені великі об’єкти.

3. Орієнтовний масштаб, в якому дані можуть використовуватись без суттєвих втрат точності та деталізації.

Результати оцінювання:

– Microsoft Building Footprints продемонстрував найкращі показники повноти: мінімальна площа об’єкта – лише 0,2542 м^2, що свідчить про фіксацію навіть найменших будівель. Максимальна площа сягає 28 983,83 м². Орієнтовний масштаб використання даних – від 1:500 до 1:50 000, що свідчить про високу адаптивність даного джерела.

– OpenStreetMap (OSM) показав вищий поріг мінімальної площі (7,55 м²), що вказує на обмежену здатність передавати малі об’єкти. Максимальна площа будівель схожа з Microsoft – 28 840,94 м². Орієнтовний масштаб використання – від 1:2000 до 1:50 000.

Google Satellite та Esri Imagery мали схожі результати: мінімальні площі становили близько 59 м², а максимальні – близько 9700 м^2. Це свідчить про те, що в обох джерелах об’єкти нижче певного порогу не були розпізнані. Орієнтовний масштаб: 1:5000 – 1:50 000.

Bing Maps показав подібний рівень повноти до Google та Esri: мінімальна площа – 45,24 м², максимальна – 9624,43 м². Орієнтовний масштаб використання також становить 1:5000 – 1:50 000.

Отже, найвищу повноту забезпечують дані Microsoft Building Footprints, оскільки охоплюють як найменші, так і найбільші об’єкти, що підтверджується широким діапазоном площ. Інші джерела (особливо супутникові) не фіксують будівлі малої площі, що обмежує їхню придатність для масштабів крупніше ніж 1:5000. OpenStreetMap хоча і охоплює великі об’єкти, має вищий поріг мінімальної площі, що також знижує рівень повноти для дрібних форм забудови.

3.4 Аналіз результатів оцінки якості відкритих геопросторових даних

У ході проведення оцінки було досліджено якість геопросторових даних з п’яти відкритих джерел (OSM, Microsoft Building Footprints, Google Satellite, Esri Imagery, Bing Maps) для кількох населених пунктів. Оцінювання охоплювало три основні критерії: детальність, повнота та позиційна точність.

Детальність:

Оцінка детальності ґрунтувалася на припущенні про те, що геометрія будівель у межах масштабу 1:2000 повинна зберігати точність у формі – зокрема, не допускати перевищення довжини сторін контурів понад 40 метрів. Аналіз виконано шляхом перетворення полігональних об’єктів у лінійні елементи та подальшого обчислення довжин їхніх сегментів. За отриманими результатами, було сформовано розподіл об’єктів на три групи: ті, що повністю відповідають критерію, частково відповідають (окремі сторони перевищують допустиме значення), та ті, що не відповідають зовсім.

За результатами порівняння встановлено, що найбільшу відповідність критерію детальності продемонстрували Google Satellite (59,1%), Bing Maps (57,7%) і Esri Imagery (56,4%). Дані Microsoft Building Footprints також показали високий рівень відповідності (56,1%), тоді як OpenStreetMap продемонстрував

лише 25,2% об’єктів, що повністю задовольняють задані параметри, що свідчить про нижчий рівень геометричної деталізації цього джерела.

Повнота:

Повнота джерела оцінювалася шляхом аналізу мінімальної та максимальної площі будівель у межах вибірки, що дозволяє оцінити здатність джерела відображати як дрібні, так і великі об’єкти. На основі отриманих значень було визначено орієнтовний масштаб, у якому дані кожного джерела можуть використовуватись із прийнятним рівнем достовірності.

Найвищі показники повноти виявлено у наборі Microsoft Building Footprints, де мінімальна площа будівлі становить 0,25 м², а максимальна – понад 28 000 м², що вказує на придатність цього джерела для масштабів від 1:500 до 1:50 000. Дані OpenStreetMap також охоплюють великий спектр об’єктів, однак мають значно вищий поріг мінімальної площі (понад 7 м²). Супутникові джерела (Google, Bing, Esri) характеризуються мінімальними площами в межах 45–60 м², що обмежує їх придатність лише для середньомасштабного картографування (від 1:5000 до 1:50 000).

Позиційна точність:

У межах даної роботи оцінити позиційну точність не вдалося через значну різницю в геометрії будівель між джерелами та еталонним набором. Зокрема, будівлі мали різну кількість вершин, спрощену або змінену конфігурацію, що не дозволило виконати коректне зіставлення координат відповідних точок.

Висновок до розділу 3: На основі проведеного аналізу можна стверджувати, що найбільш повні й деталізовані просторові характеристики об’єктів містить набір Microsoft Building Footprints, що робить його оптимальним для задач, що потребують фіксації будівель різного розміру. Супутникові сервіси (Google, Bing, Esri) забезпечують кращу деталізацію великих об’єктів, проте не охоплюють малі форми забудови, що обмежує їхнє застосування в масштабах, дрібніших за 1:5000. Дані OpenStreetMap мають високу варіативність: попри відносну повноту охоплення, детальність часто не відповідає вимогам для високоточного картографування.

ВИСНОВОК

У межах дипломної роботи було досліджено питання оцінки якості відкритих геопросторових даних, які все частіше використовуються в геоінформаційних дослідженнях. Робота поєднувала як загальнотеоретичні підходи до оцінювання якості, так і практичне застосування методів аналізу в середовищі ArcGIS.

Під час дослідження було розроблено структуру бази геоданих, що дозволила впорядкувати просторову інформацію та забезпечити зручність роботи з нею. Також визначено критерії оцінки, серед яких повнота, детальність і позиційна точність. Особлива увага приділялася саме першим двом показникам через специфіку початкових даних. У рамках методики було реалізовано геометричну трансформацію об’єктів, обчислення довжин та площ, подальше групування отриманих результатів і їхню візуалізацію для полегшення аналізу.

За результатами аналізу встановлено, що найбільш повним за охопленням є набір Microsoft Building Footprints, що охоплює значну кількість як великих, так і малих будівель. Найвищу детальність продемонстрували супутникові сервіси Google, Bing та Esri, тоді як OpenStreetMap, попри свою відкритість і оновлюваність, поступається за критерієм геометричної точності. Оцінка позиційної точності не проводилася через значні відмінності в конфігурації будівель між джерелами.

Таким чином, можна зробити висновок, що комплексна оцінка якості відкритих геопросторових даних дозволяє аргументовано підходити до вибору джерел для просторового аналізу. Запропонована методика має потенціал для застосування до інших територій та наборів даних у майбутніх дослідженнях.

СПИСОК ВИКОРИСТАНИХ ДЖЕРЕЛ

1. України, З. (2020). Про національну інфраструктуру геопросторових даних. (прийнято ВР 13.04. 2020 р., вводиться в дію з 1.01. 2021 р.)[Електронний ресурс].– Режим доступу: https://zakon.rada.gov.ua/laws/show/554-20#Text .

2. Мінагрополітики. (2021). Технічні вимоги та методи забезпечення інтероперабельності і сумісності наборів геопросторових даних та геоінформаційних сервісів (№ z0024-22; наказ від 10.11.2021 № 347) [Електронний ресурс]. – Режим доступу: https://zakon.rada.gov.ua/laws/show/z0024-22#Text .

3. Мінагрополітики. (2021). Технічні вимоги до метаданих для наборів геопросторових даних і геоінформаційних сервісів національної інфраструктури геопросторових даних (№ z0021-22; наказ від 10.11.2021 № 347)[Електронний ресурс]. – Режим доступу: https://zakon.rada.gov.ua/laws/show/z0021-22#n18 .

4. Кабінет Міністрів України. (2021). Про затвердження Порядку функціонування національної інфраструктури геопросторових даних (постанова від 26.05.2021 № 532-2021-п) [Електронний ресурс]. – Режим доступу: https://zakon.rada.gov.ua/laws/show/532-2021-п#Text .

5. Європейський парламент і Рада ЄС. (2007). Директива 2007/2/ЄС про створення інфраструктури просторової інформації у Європейському Співтоваристві (INSPIRE) від 14.03.2007 р. [Електронний ресурс]. – Режим доступу: https://zakon.rada.gov.ua/laws/show/984_002-07#Text .

6. ISO. (2013). ISO 19157:2013 – Geographic information – Data quality. [Електронний ресурс]. – Режим доступу: https://www.iso.org/ru/standard/32575.html .

7. ISO. (2014). ISO 19115-1:2014 – Geographic information — Metadata. [Електронний ресурс]. – Режим доступу: https://www.iso.org/ru/standard/53798.html .

8. Open Geospatial Consortium. (2023). OGC Standards and Publications. [Електронний ресурс]. – Режим доступу: https://www.ogc.org/publications/ .

9. Caspary, W. (1993). Qualitätsaspekte bei Geo-Informationssystemen. *ZfV*, *118*(8), 9.

10. Ostensen, O., ìPersonal communication with the chairman of the ISO / TC 211 during the 31st Digital Geographic Information Working Group (DGIWG) Steering Committee Meeting in Oslo on May 3-4î, Norway, (1999).

11. ISO. (1999). ISO 2859-1:1999 – Sampling procedures for inspection by attributes. [Електронний ресурс]. – Режим доступу: https://www.iso.org/ru/standard/1141.html .

12. ISO. (2022). ISO 3951-1:2022 – Sampling procedures for inspection by variables. [Електронний ресурс]. – Режим доступу: https://www.iso.org/ru/standard/74706.html .

13. Міністерство аграрної політики та продовольства України. (2024). ПОРЯДОК топографічної зйомки у масштабах 1:5000, 1:2000, 1:1000 та 1:500 (№21-6010-05.1/20243 від 31.07.2024)[Електронний ресурс]. – Режим доступу: https://land.gov.ua/wp-content/uploads/2024/08/poryadok.pdf .

14. Xia, J. (2012). Metrics to measure open geospatial data quality. Issues in Science and Technology Librarianship, 68. [Електронний ресурс]. – Режим доступу: http://www.istl.org/12-winter/article1.html?a_aid=3598aabf .

15. Taştan, H., & Altan, M. O. (2011). Spatial Data Quality. Proceedings of 3rd Turkish-German Joint Geodetic Days. [Електронний ресурс]. – Режим доступу: https://surl.lu/yiuwam .

16. Dawod, G. M., & Ascoura, I. E. (2021). The validity of open-source elevations for different topographic map scales and geomatics applications. Journal of Geographic Information System, 13(02), 148. [Електронний ресурс]. – Режим доступу: https://www.scirp.org/pdf/jgis_2021031811280911.pdf .

17. Moradi, M., Roche, S., & Mostafavi, M. A. (2023). Evaluating osm building footprint data quality in québec province, canada from 2018 to 2023: A comparative study. *Geomatics*, *3*(4), 541-562. [Електронний ресурс]. – Режим доступу: https://www.mdpi.com/2673-7418/3/4/29# .