

МІНІСТЕРСТВО ОСВІТИ І НАУКИ УКРАЇНИ
ПОЛІСЬКИЙ НАЦІОНАЛЬНИЙ УНІВЕРСИТЕТ

Факультет інформаційних технологій, обліку та фінансів
Кафедра комп'ютерних технологій
і моделювання систем

Кваліфікаційна робота
на правах рукопису

Бойко Максим Володимирович

УДК 504.064:004.9:528.85

КВАЛІФІКАЦІЙНА РОБОТА

Виявлення рівня забрудненості атмосфери за даними мультиспектрального
сканеру Terra

(тема роботи)

122 «Комп'ютерні науки»

(шифр і назва спеціальності)

Подається на здобуття освітнього ступеня бакалавр

кваліфікаційна робота містить результати власних досліджень. Використання
ідей, результатів і текстів інших авторів мають посилання на відповідне
джерело

(підпис, ініціали та прізвище здобувача вищої освіти)

Керівник роботи
Топольницький Павло Петрович
(прізвище, ім'я, по батькові)
кандидат технічних наук, доцент
(науковий ступінь, вчене звання)

Житомир – 2025

Висновок кафедри _____
за результатами попереднього захисту: _____

Протокол засідання кафедри _____
№ _____ від « _____ » _____ 20 _____ р.

Завідувач кафедри _____

(науковий ступінь, вчене звання)

(підпис)

(прізвище, ім'я, по батькові)

« _____ » _____ 20 _____ р.

Результати захисту кваліфікаційної роботи

Здобувач вищої освіти Бойко Максим Володимирович захистив (ла)

(прізвище, ім'я, по батькові)

кваліфікаційну роботу з оцінкою:

сума балів за 100-бальною шкалою _____

за шкалою ECTS _____

за національною шкалою _____

Секретар ЕК

(науковий ступінь, вчене звання)

(підпис)

(прізвище, ім'я, по батькові)

АНОТАЦІЯ

Бойко М. В. Виявлення рівня забрудненості атмосфери за даними мультиспектрального сканеру Terra – Кваліфікаційна робота на правах рукопису.

Кваліфікаційна робота на здобуття освітнього ступеня бакалавра за спеціальністю 122 – Комп’ютерні науки. – Поліський національний університет, Житомир, 2025.

У роботі розглянуто питання автоматизації екологічного моніторингу на основі супутникових мультиспектральних даних та сучасних інформаційних технологій. Проведено аналіз предметної області, описано теоретичні основи дистанційного зондування Землі, методи штучного інтелекту для аналізу екологічних даних і алгоритми обробки мультиспектральної інформації. Запропоновано архітектуру програмної системи, що інтегрується з зовнішніми сервісами (Copernicus, Sentinel Hub, метеорологічні API) для збору, обробки та візуалізації екологічних показників. Описано структуру бази даних, технологічний стек, інструменти розробки та модулі системи. Система забезпечує автоматизований збір, аналіз і представлення даних про стан атмосфери, що підвищує ефективність екологічного моніторингу та може бути основою для впровадження інтелектуальних алгоритмів прогнозування.

Ключові слова: екологічний моніторинг, мультиспектральні супутникові дані, сканер Terra, забруднення атмосфери, автоматизація збору даних.

SUMMARY

Boyko M. V. Detection of atmospheric pollution level using Terra multispectral scanner data – Qualification work in the form of a manuscript.

Qualification work for obtaining a bachelor's degree in specialty 122 – Computer Science. – Polissia National University, Zhytomyr, 2025.

The paper considers the issue of automation of environmental monitoring based on satellite multispectral data and modern information technologies. The subject area is analyzed, the theoretical foundations of remote sensing of the Earth are described, artificial intelligence methods for analyzing environmental data and algorithms for processing multispectral information are described. The architecture of a software system that integrates with external services (Copernicus, Sentinel Hub, meteorological APIs) for collecting, processing and visualizing environmental indicators is proposed. The database structure, technological stack, development tools and system modules are described. The system provides automated collection, analysis and presentation of data on the state of the atmosphere, which increases the efficiency of environmental monitoring and can be the basis for the implementation of intelligent forecasting algorithms.

Keywords: environmental monitoring, multispectral satellite data, Terra scanner, atmospheric pollution, automation of data collection.

ЗМІСТ

ПЕРЕЛІК УМОВНИХ ПОЗНАЧЕНЬ	6
ВСТУП.....	7
РОЗДІЛ 1. АНАЛІЗ ПРЕДМЕТНОЇ ОБЛАСТІ ТА ТЕОРЕТИЧНІ ОСНОВИ ЕКОЛОГІЧНОГО МОНІТОРИНГУ	9
1.1 Технології дистанційного зондування Землі.....	9
1.2 Методи штучного інтелекту для аналізу екологічних даних.....	10
1.3 Алгоритми обробки мультиспектральних даних	11
Висновки до першого розділу	13
РОЗДІЛ 2. МЕТОДОЛОГІЯ ТА АЛГОРИТМИ ОБРОБКИ ЕКОЛОГІЧНИХ ДАНИХ	14
2.1 Системний аналіз та вимоги	14
2.2 Архітектура системи	17
2.3 Розроблена бази даних	18
Висновки до другого розділу	20
РОЗДІЛ 3. ПРОЕКТУВАННЯ ТА АРХІТЕКТУРА СИСТЕМИ.....	22
3.1 Інтеграція з зовнішніми сервісами	22
3.2 Технологічний стек та інструменти розробки, модуль збору та обробки даних	23
Висновки до третього розділу	25
ВИСНОВКИ	26
СПИСОК ВИКОРИСТАНИХ ДЖЕРЕЛ	27
ДОДАТКИ.....	32

ПЕРЕЛІК УМОВНИХ ПОЗНАЧЕНЬ

БД – база даних;

ГІС – геоінформаційні системи;

ІС – інформаційна система;

МН – машинне навчання;

ПЗ – програмне забезпечення;

ШІ – штучний інтелект;

API – application programming interface;

UML – unified modeling language.

ВСТУП

У сучасному світі проблема забруднення атмосфери набуває особливої гостроти через інтенсивну індустріалізацію, урбанізацію та зростання транспортного навантаження. За даними Всесвітньої організації охорони здоров'я, забруднення повітря щорічно призводить до смерті понад 7 мільйонів людей у світі та є одним із найбільших екологічних ризиків для здоров'я людини [1, 2]. В Україні ситуація ускладнюється через високу концентрацію промислових підприємств, застарілі технології та недостатню ефективність існуючих систем екологічного моніторингу [3, 4, 5].

Традиційні методи моніторингу якості атмосферного повітря, що базуються на мережі стаціонарних станцій, мають обмежену просторову роздільну здатність та не забезпечують повного покриття території. Водночас стрімкий розвиток технологій дистанційного зондування Землі, зокрема супутникових систем NASA Terra, Aqua та ESA Sentinel-5P, відкриває нові можливості для створення комплексних систем екологічного моніторингу з високою просторово-часовою роздільною здатністю. Особливого значення набуває інтеграція сучасних інформаційних технологій, включаючи геоінформаційні системи (ГІС), хмарні обчислення та алгоритми штучного інтелекту, для обробки великих обсягів супутникових даних та створення інтелектуальних систем підтримки прийняття рішень у сфері екологічного управління.

За темою кваліфікаційної роботи було також опубліковано тези на таких конференціях:

- Міжнародна науково-практична конференція «Envolving science: Theories, Discoveries and Practical Outcomes».
- Збірник праць учасників всеукраїнської науково-практичної конференції здобувачів вищої освіти і молодих вчених «Інформаційні технології та моделювання систем»

Мета дослідження полягає в розробці інтегрованої геоінформаційної системи для моніторингу екологічного стану атмосфери з використанням даних мультиспектрального сканера Terra та алгоритмів машинного навчання, що забезпечить підвищення ефективності екологічного контролю та якості прийняття управлінських рішень.

Об'єкт дослідження: процеси моніторингу та аналізу екологічного стану атмосфери з використанням технологій дистанційного зондування Землі.

Предмет дослідження: методи, алгоритми та програмні засоби створення геоінформаційних систем для обробки даних мультиспектрального сканування з метою визначення та прогнозування рівнів забруднення атмосфери.

Методи дослідження:

- системний аналіз
- методи цифрової обробки зображень
- статистичний аналіз
- методи машинного навчання
- методи просторового аналізу

РОЗДІЛ 1. АНАЛІЗ ПРЕДМЕТНОЇ ОБЛАСТІ ТА ТЕОРЕТИЧНІ ОСНОВИ ЕКОЛОГІЧНОГО МОНІТОРИНГУ

1.1 Технології дистанційного зондування Землі

Супутникові системи дистанційного зондування Землі відіграють ключову роль у сучасному екологічному моніторингу атмосфери. Проєкт NASA Earth Observing System (EOS) включає супутники Terra і Aqua, які розташовані на сонячно-синхронній орбіті та оснащені спектрорадіометрами середньої роздільної здатності (MODIS). Вони забезпечують регулярне отримання мультиспектральних даних про стан атмосфери, поверхні суші та океанів, а також дозволяють відстежувати аерозолі, парникові гази і хмарність у глобальному масштабі [7].

Sentinel-5P – це спеціалізований супутник Європейського космічного агентства, призначений для моніторингу якості повітря і складу атмосфери. Оснащений інструментом TROPOMI, він забезпечує високу просторову (3.5×5.5 км) і часову роздільну здатність, що дозволяє щоденно відстежувати концентрації озону, діоксиду азоту, чадного газу, метану, формальдегіду та інших забруднювачів. Sentinel-5P ефективно виявляє джерела забруднення у міських і промислових зонах, що робить його незамінним інструментом для екологічного моніторингу та оцінки впливу на здоров'я населення [7, 8].

Мультиспектральні дані, які отримують супутники Terra, Aqua і Sentinel-2, охоплюють кілька широких діапазонів електромагнітного спектра (видимий, ближній інфрачервоний, короткохвильовий інфрачервоний), що дозволяє аналізувати різноманітні характеристики земної поверхні та атмосфери. Sentinel-2, зокрема, забезпечує високоякісні знімки з просторовою роздільною здатністю від 10 до 60 м, що широко використовується для моніторингу рослинності, водних ресурсів і змін у землекористуванні.

Гіперспектральні дані, на відміну від мультиспектральних, містять сотні вузькоспектральних каналів, що дозволяє більш детально і точно ідентифікувати хімічний склад атмосфери, типи рослинності або забруднювачів. Такі дані застосовуються у складних екологічних

дослідженнях, зокрема для виявлення забруднень, аналізу стану екосистем і прогнозування кліматичних змін.

Обробка супутникових даних включає кілька етапів: калібрування і корекція знімків, фільтрація шумів і помилок, геоприв'язка, а також класифікація і аналіз спектральних характеристик. Для моніторингу атмосфери застосовуються спеціалізовані алгоритми, які дозволяють виділяти концентрації окремих газів і аерозолів, наприклад, діоксиду азоту (NO_2), чадного газу (CO), формальдегіду (HCHO) та озону (O_3) [9].

Особливу увагу приділяють виявленню і усуненню статистично ненадійних даних, що можуть виникати через хмарність, відблиски або технічні помилки. Автоматизовані системи обробки, такі як розроблені для Sentinel-5P, забезпечують оперативний аналіз і формування карт забруднення повітря з деталізацією до міських районів, що є важливим для прийняття управлінських рішень у сфері екології [9].

1.2 Методи штучного інтелекту для аналізу екологічних даних

Машинне навчання (МН) є одним із ключових методів штучного інтелекту, що активно застосовується для аналізу екологічних даних. Завдяки здатності обробляти великі обсяги інформації та виявляти приховані закономірності, МН дозволяє ефективно моніторити стан довкілля, прогнозувати зміни клімату і оцінювати екологічні ризики. Використання алгоритмів машинного навчання допомагає автоматизувати аналіз даних із супутникових знімків, сенсорів якості повітря та інших джерел, що значно підвищує точність і швидкість прийняття рішень у сфері охорони довкілля [10].

Для обробки екологічних даних широко застосовуються алгоритми класифікації та регресії. Класифікаційні моделі використовуються для розпізнавання типів забруднень, класифікації стану рослинності, визначення категорій якості повітря або води. Регресійні алгоритми дозволяють

прогнозувати кількісні показники, такі як концентрації шкідливих газів, рівень забруднення або зміну кліматичних параметрів у часі. Ці методи допомагають створювати точні моделі, які враховують складні взаємозв'язки між різними екологічними факторами [11; 12].

Глибокі нейронні мережі, зокрема згорткові нейронні мережі (CNN), ефективно застосовуються для аналізу супутникових і аерофотознімків, що використовуються в екологічному моніторингу. Вони дозволяють автоматично виявляти зміни в землекористуванні, поширення забруднень, стан лісів і водних об'єктів. Завдяки здатності до навчання на великих обсягах зображень, нейронні мережі забезпечують високу точність класифікації і сегментації екологічних об'єктів, що значно покращує якість моніторингу і прийняття управлінських рішень [10].

Прогнозування екологічних показників за допомогою штучного інтелекту базується на аналізі історичних даних і виявленні трендів у зміні параметрів довкілля. Алгоритми машинного навчання дозволяють створювати моделі, які передбачають екстремальні погодні явища, рівень забруднення повітря, зміни клімату та інші критичні показники. Це дає змогу органам влади і підприємствам оперативно реагувати на потенційні загрози, планувати заходи з адаптації та мінімізації негативного впливу на навколишнє середовище [13; 14].

1.3 Алгоритми обробки мультиспектральних даних

Існуючі алгоритми обробки мультиспектральних супутникових даних охоплюють комплекс методів статистичного аналізу, калібрування, корекції та візуалізації, що забезпечують точний і надійний аналіз екологічних показників.

Для кількісної оцінки просторових змінних із NetCDF-файлів застосовуються стандартні статистичні методи, які включають обчислення середнього, медіани, мінімуму, максимуму, стандартного відхилення та квантилів (25% і 75%). Перед розрахунками виконується фільтрація

пропущених або некоректних значень (NaN), що підвищує якість і коректність аналізу [15]. Ці статистичні характеристики дозволяють виявити розподіл концентрацій забруднювачів і просторову неоднорідність у досліджуваних регіонах.

Для 2D та 3D даних широко застосовують побудову кольорових карт (pcolormesh) із кастомним колормепом, що відображає просторовий розподіл концентрацій атмосферних компонентів, таких як NO_2 , O_3 , CO , SO_2 , HCHO , CH_4 та аерозолів [16]. У випадках аналізу одномірних даних або при виникненні помилок використовують гістограми розподілу значень з додаванням оцінки густини розподілу (kernel density estimation), що покращує інтерпретацію даних [15].

Існують автоматизовані алгоритми, які здійснюють пошук пріоритетних змінних, характерних для мультиспектральних супутникових даних. Зокрема, це концентрації основних забруднювачів атмосфери (NO_2 , O_3 , CO , SO_2 , HCHO , CH_4) та аерозолів, що дозволяє фокусувати аналіз на найбільш інформативних параметрах для екологічного моніторингу [17].

Для роботи з багатовимірними супутниковими даними використовують бібліотеки, оптимізовані для ефективного зчитування, обробки і аналізу великих масивів даних у форматі NetCDF, зокрема `xarray`. Підтримується пакетна обробка файлів, що дозволяє автоматизувати аналіз великих обсягів супутникових даних, наприклад, отриманих із платформи WEkEO [17].

Числові розрахунки виконуються за допомогою бібліотеки `numpy`, що забезпечує високу продуктивність і точність обчислень. Для оцінки густини розподілу застосовується модуль `scipy.stats`, який реалізує методи kernel density estimation. Візуалізація результатів здійснюється за допомогою `matplotlib`, що дозволяє створювати гнучкі і інформативні графіки та карти.

Висновки до першого розділу

У ході дослідження було проаналізовано основні технології дистанційного зондування Землі, які використовуються для екологічного моніторингу атмосфери. Розглянуто супутникові системи NASA Earth Observing System, зокрема Terra і Aqua, що забезпечують регулярне отримання мультиспектральних даних про стан атмосфери, поверхні суші та океанів, а також моніторинг аерозолів і парникових газів. Окремо висвітлено можливості супутника Sentinel-5P Європейського космічного агентства. Описано особливості мультиспектральних та гіперспектральних даних, які дають змогу детально аналізувати хімічний склад атмосфери, типи рослинності та забруднення. Також розглянуто етапи обробки супутникових даних, включно з калібруванням, атмосферною корекцією, фільтрацією шумів і класифікацією спектральних характеристик, що забезпечує високу точність моніторингу.

У роботі також висвітлено застосування методів штучного інтелекту, зокрема машинного навчання, алгоритмів класифікації, регресії та глибоких нейронних мереж для аналізу екологічних даних, автоматичного виявлення змін у землекористуванні та прогнозування екологічних показників. Застосування цих методів підвищує швидкість і точність обробки великих обсягів інформації.

РОЗДІЛ 2. МЕТОДОЛОГІЯ ТА АЛГОРИТМИ ОБРОБКИ ЕКОЛОГІЧНИХ ДАНИХ

2.1 Системний аналіз та вимоги

Системний аналіз інформаційної системи починається з визначення функціональних та нефункціональних вимог, які формують основу для подальшої розробки. Функціональні вимоги описують конкретні можливості, які система має надавати користувачам. Зокрема, це можливість перегляду списку територій для моніторингу, вибору конкретної території, налаштування параметрів моніторингу, перегляду карт забруднення та результатів аналізу, а також експорту отриманих даних. Важливою функцією є також можливість користувача надсилати запити на додавання нових територій, а для адміністратора – керувати територіями, кешем та отримувати дані з зовнішніх джерел, таких як Sentinel API.

Нефункціональні вимоги визначають певні якісні характеристики системи. Серед них – продуктивність, що забезпечує швидку обробку великих обсягів супутникових даних, масштабованість, що дозволяє розширювати систему без втрати ефективності; надійність і безпека, які гарантують збереження даних і обмеження доступу до адміністративних функцій, а також зручність користування, що передбачає інтуїтивний інтерфейс і зрозумілу візуалізацію інформації. Крім того, система повинна підтримувати інтеграцію з зовнішніми API для автоматичного оновлення даних.

Сценарії використання, відображені на діаграмі прецедентів (рис. 2.1), ілюструють взаємодію користувачів і системи. Користувачі виконують дії з перегляду і вибору територій, налаштування параметрів, аналізу та експорту даних, а також можуть ініціювати запити на розширення системи. Адміністратори відповідають за управління територіями і кешем, що забезпечує стабільність і актуальність інформації. Система автоматично отримує і обробляє дані з зовнішніх джерел, підтримуючи безперервний моніторинг.



Рисунок 2.1 – Uml-діаграма прецедентів

Послідовність роботи системи, відображена на діаграмі (рис. 2.2), починається з того, що користувач через веб-інтерфейс обирає територію зі списку і задає параметр для аналізу, наприклад, CO₂ чи NO₂. Далі формується запит на отримання даних про забруднення, який надходить у backend. Backend спочатку перевіряє наявність потрібної інформації у локальному кеші: якщо дані вже є, вони одразу повертаються користувачу, що прискорює роботу. Якщо ж даних у кеші немає, система ініціює запит на отримання даних для обраної території, який надходить до модуля обробки даних (Processing API).

Модуль обробки займається аналізом растрових даних, обчисленням статистики і повертає оброблені результати у backend. Після цього відбувається збереження результатів у базі даних та кеші, а також, за необхідності, виконуються додаткові обчислення. Далі результати відправляються на фронтенд, де відбувається рендеринг карти, відображення статистики, легенди та показ користувачу всіх результатів аналізу.

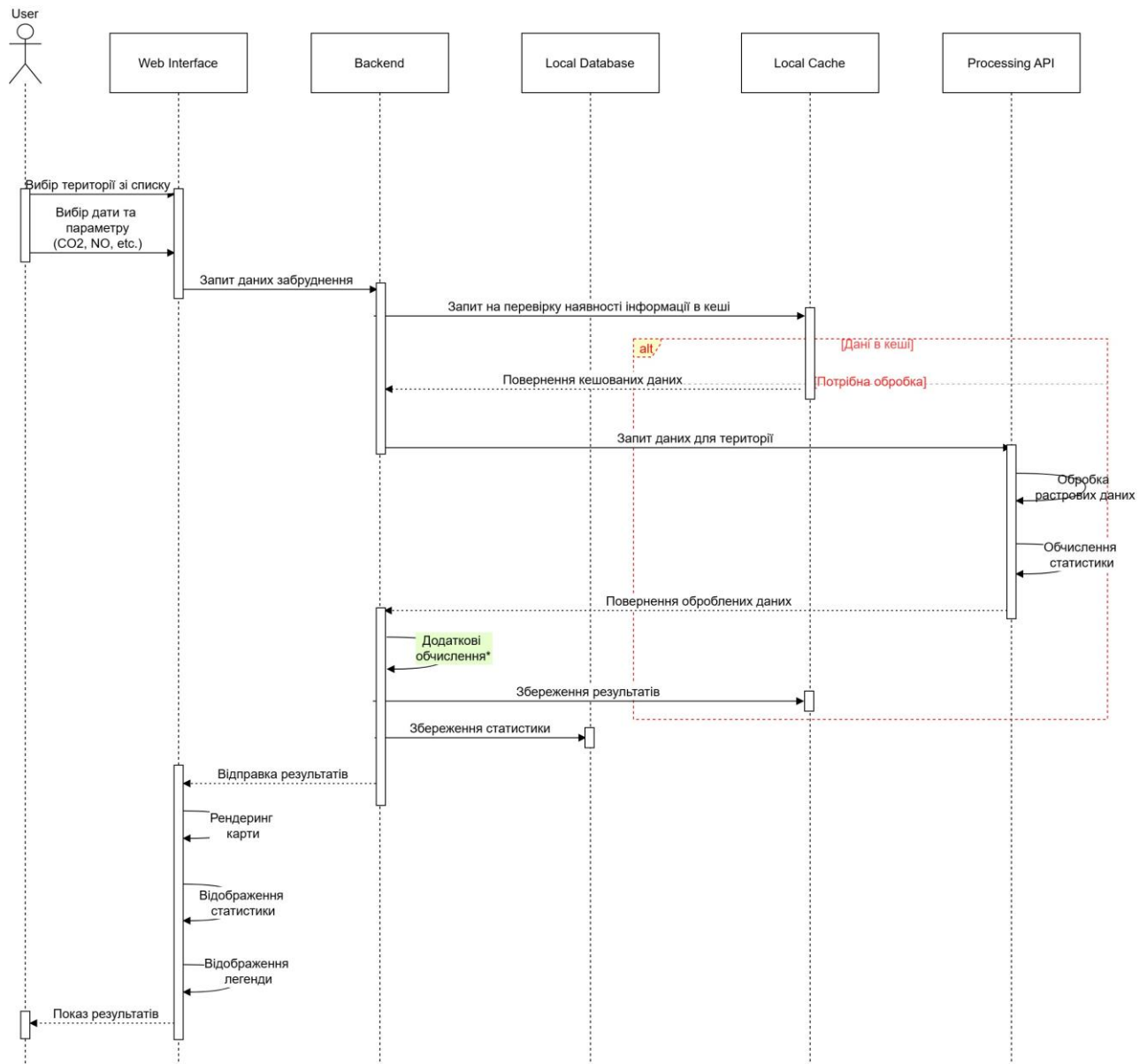


Рисунок 2.2 – UML-діаграма послідовностей

Розширення цієї послідовності показане на другій діаграмі (див. рис. A1), яка деталізує процес додаткових обчислень. Backend надсилає запит на додаткові обчислення до Processing API, а той, у свою чергу, звертається до Calculation Module. Тут відбувається кілька послідовних етапів: спочатку дані нормалізуються і приводяться до стандартного діапазону, потім розраховуються базові статистичні показники – середнє значення, медіана, мінімум, максимум, стандартне відхилення. Далі результати порівнюються з гранично допустимими концентраціями, визначається відсоток перевищення норми по території, за можливості – порівнюється з історичними даними,

розраховується динаміка змін, наприклад, зміна від попереднього місяця чи тренд за певний період.

Після цього система виконує класифікацію рівнів забруднення (низький, середній, високий), генерує текстовий опис і автоматично формує висновки про стан повітря. Усі ці результати повертаються назад у backend, а потім – користувачу. Така послідовність забезпечує не лише базовий аналіз, а й глибоку аналітику з автоматичними висновками, що робить систему зручною та інформативною для кінцевого користувача.

2.2 Архітектура системи

Архітектура системи побудована за принципами багаторівневої (layered) архітектури з чітким розділенням на компоненти, що забезпечує модульність, гнучкість і масштабованість. Основними складовими є API-сервіс, реалізований на Flask, який виступає як основний веб-сервер для прийому та обробки запитів, взаємодії з базою даних і іншими сервісами. Окремо виділено сервіс обробки даних (processing.py), який відповідає за аналіз, статистику та візуалізацію мультиспектральних супутникових даних, що дозволяє ізолювати ресурсоємні обчислення від основного API. Для інтеграції із зовнішніми джерелами інформації використовується клієнт до Copernicus STAC API (copernicus_client.py), який інкапсулює логіку роботи з зовнішніми API. Зберігання даних, таких як токени доступу, результати аналізу та історія запитів, здійснюється у базі даних SQLite з використанням ORM SQLAlchemy, що спрощує роботу з даними та їхню організацію.

Незважаючи на те, що поточна реалізація системи є монолітною, вона має чітку модульну структуру, що полегшує подальший перехід до мікросервісної архітектури. Для цього передбачено можливість винесення окремих компонентів – обробки даних, API-клієнта, роботи з базою даних – у самостійні сервіси, які можна розгортати у контейнерах (наприклад, Docker). Взаємодія між сервісами у такій архітектурі може бути організована

асинхронно з використанням HTTP, gRPC або систем черг повідомлень, що підвищить масштабованість і стійкість системи.

Масштабування системи передбачає запуск кількох інстансів сервісів обробки даних і API, що дозволяє розподілити навантаження і підвищити продуктивність. Відмовостійкість досягається за рахунок дублювання сервісів, використання зовнішньої бази даних із резервним копіюванням, а також моніторингу стану компонентів. Для повноцінної відмовостійкості рекомендується застосовувати оркестратори контейнерів, такі як Kubernetes або Docker Swarm, а також інтегрувати зовнішні сервіси зберігання даних.

2.3 Розроблена бази даних

Для реалізації системи екологічного моніторингу була розроблена реляційна база даних, що забезпечує ефективне зберігання, обробку та аналіз екологічних даних із різних джерел. База даних (рис. 2.3) спроектована з урахуванням специфіки роботи з геопросторовими даними, великими обсягами вимірювань та необхідності підтримки багатокористувацького режиму. Центральною таблицею системи є `environmental_data`, яка зберігає всі екологічні вимірювання з часовими мітками, географічними координатами та метаданими. Ця таблиця пов'язана з довідниками джерел даних (`data_sources`), географічних регіонів (`geographic_regions`) та типів параметрів (`parameter_types`), що забезпечує нормалізацію даних та гнучкість системи.

Таблиця `users` реалізує систему аутентифікації та авторизації з підтримкою різних ролей користувачів, включаючи адміністраторів, аналітиків та звичайних користувачів. Для забезпечення безпеки паролі зберігаються у хешованому вигляді, а система підтримує відстеження активності користувачів.

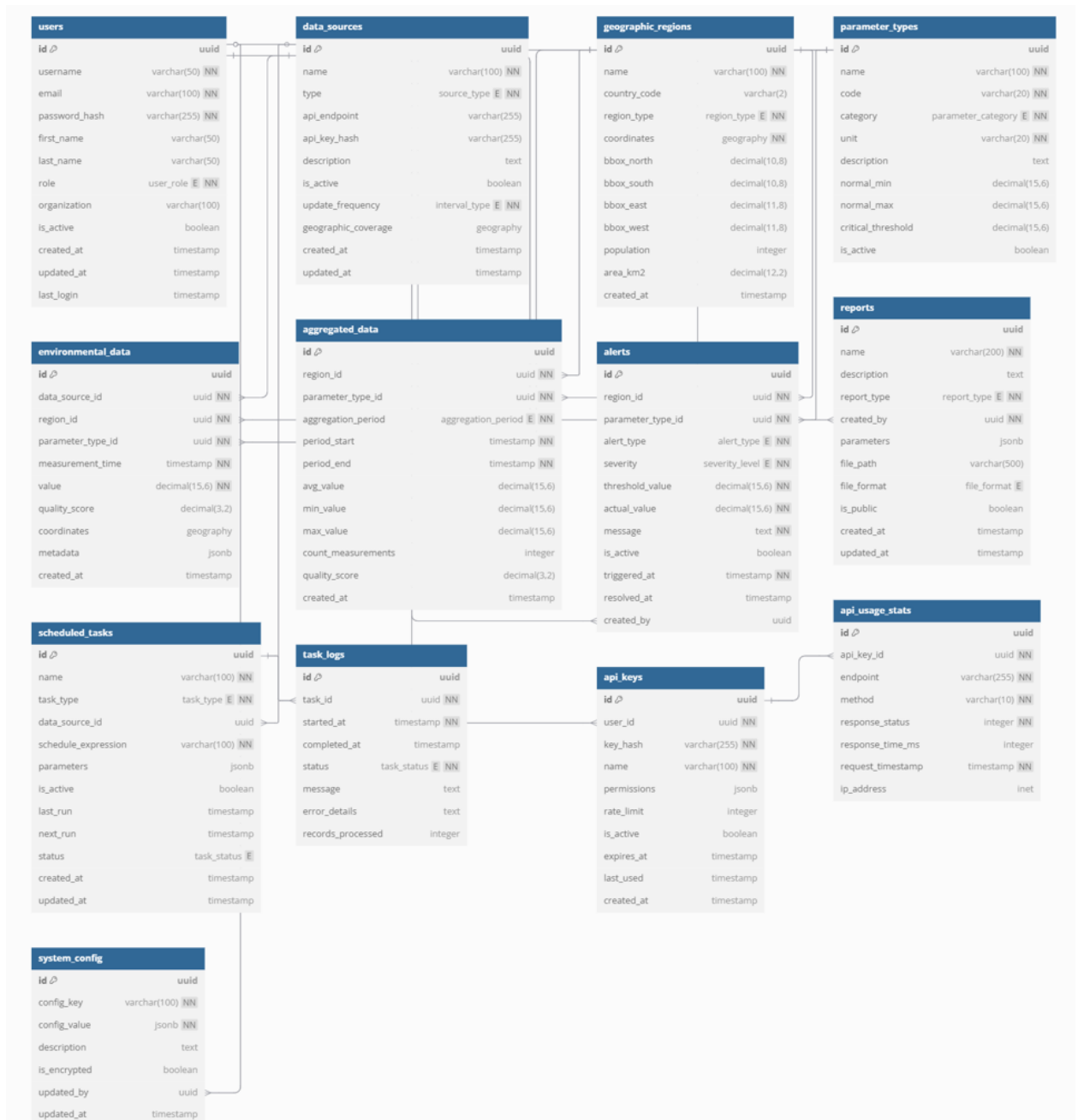


Рисунок 2.3 – Схема бази даних системи

Для оптимізації продуктивності при роботі з великими обсягами даних створена таблиця `aggregated_data`, яка містить попередньо розраховані статистичні показники за різні періоди часу. Це дозволяє швидко отримувати аналітичні дані без необхідності обробки мільйонів записів у режимі реального часу.

Система моніторингу та сповіщень реалізована через таблицю `alerts`, яка відстежує перевищення екологічних норм та генерує попередження різного

рівня критичності. Алерти пов'язані з конкретними регіонами та типами параметрів, що дозволяє забезпечити точну локалізацію проблем.

Автоматизація збору та обробки даних забезпечується через таблиці `scheduled_tasks` та `task_logs`, які реалізують планувальник завдань з детальним логуванням виконання. Це дозволяє системі автономно отримувати дані з різних джерел, обробляти їх та генерувати звіти за заданим розкладом.

Для інтеграції з зовнішніми системами створена таблиця `api_keys` з відповідною системою контролю доступу та моніторингу використання через `api_usage_stats`. Це забезпечує безпечний доступ до даних через API та дозволяє відстежувати навантаження на систему.

Гнучкість конфігурації системи забезпечується через таблицю `system_config`, яка дозволяє зберігати параметри у форматі JSON з підтримкою шифрування чутливих даних. Таблиця `reports` надає можливість зберігання та управління згенерованими звітами різних типів та форматів.

Для забезпечення високої продуктивності створені індекси на найбільш часто використовувані поля, включаючи геопросторові індекси для швидкого пошуку за координатами. Індекси побудовані з урахуванням типових запитів системи, таких як вибірка даних за часовим періодом, географічним регіоном та типом параметра.

Розроблена структура бази даних забезпечує масштабованість системи, можливість горизонтального розподілу даних за часовими періодами та географічними регіонами, а також підтримку резервного копіювання та відновлення даних. Використання UUID як первинних ключів забезпечує можливість розподіленого зберігання даних та їх реплікації між різними інстансами системи.

Висновки до другого розділу

Проведено аналіз функціональних та нефункціональних вимог системи на основі чого розроблено технічне завдання. Описана архітектура системи

забезпечує автоматизований збір, обробку, аналіз і візуалізацію екологічних даних із зовнішніх джерел, дозволяє легко масштабувати й розширювати функціонал, інтегрувати нові сервіси та підтримувати сучасний рівень надійності й гнучкості. Наведено таблицю бази даних системи, яка містить основні дані необхідні для функціонування системи.

РОЗДІЛ 3. ПРОЕКТУВАННЯ ТА АРХІТЕКТУРА СИСТЕМИ

3.1 Інтеграція з зовнішніми сервісами

Система інтегрується з низкою зовнішніх сервісів для забезпечення отримання та обробки мультиспектральних супутникових даних, а також додаткової екологічної інформації. Основним джерелом даних є Copernicus API, зокрема Sentinel Hub, через який реалізовано окремий клієнт для формування запитів до STAC API. Цей клієнт відповідає за отримання даних про стан атмосфери, концентрації забруднювачів і аерозолів, а також їх подальшу передачу в систему для аналізу.

Код реалізації виглядає наступним чином:

```
import requests

class CopernicusClient:
    def __init__(self):
        self.base_url =
"https://catalogue.dataspace.copernicus.eu/stac/collections/SENTINEL-5P/items"

    def fetch_data(self, bbox, date):
        # bbox: [minLon, minLat, maxLon, maxLat]
        # date: 'YYYY-MM-DD' (UTC)
        params = {
            "bbox": f"{bbox[0]},{bbox[1]},{bbox[2]},{bbox[3]}",
            "datetime": f"{date}T00:00:00Z/{date}T23:59:59Z",
            "limit": 10
        }
        response = requests.get(self.base_url, params=params)
        if response.status_code == 200:
            return response.json()
        else:
            return None
```

Для розширення джерел екологічних даних передбачена можливість підключення до NASA Earthdata API. Цей сервіс надає широкий спектр

супутникових продуктів, включно з кліматичними та атмосферними показниками, що підвищує точність і повноту моніторингу. NASA Earthdata Search API дозволяє здійснювати пошук, візуалізацію та завантаження даних із великої кількості колекцій, що охоплюють різні аспекти стану Землі.

Крім того, система може інтегруватися з метеорологічними сервісами, такими як OpenWeatherMap або WeatherAPI, для врахування погодних умов – температури, вологості, швидкості і напрямку вітру тощо. Це дозволяє підвищити якість аналізу екологічної ситуації, враховуючи вплив метеоумов на розповсюдження забруднювачів.

Для оперативного інформування користувачів про критичні рівні забруднення або інші важливі події реалізовано модуль сповіщень. Він підтримує надсилання повідомлень через email, SMS, push-нотифікації або інтеграцію з популярними месенджерами, що забезпечує своєчасне реагування на екологічні загрози.

3.2 Технологічний стек та інструменти розробки, модуль збору та обробки даних

Технологічний стек та інструменти розробки базуються на мові програмування Python, яка є основою для бекенд-розробки, обробки даних та інтеграції з зовнішніми API. Для створення веб-сервісу використовується фреймворк Flask, що дозволяє реалізувати REST API та веб-інтерфейс. Робота з базою даних організована через ORM SQLAlchemy, що спрощує взаємодію з реляційною базою SQLite, а керування міграціями бази здійснюється за допомогою Flask-Migrate. Для обробки і аналізу багатовимірних наукових даних застосовуються бібліотеки numpy та scipy, які забезпечують ефективні інструменти для математичних і статистичних операцій. Візуалізація результатів реалізується за допомогою matplotlib, що дозволяє створювати інформативні графіки та карти. Інтеграція з зовнішніми API здійснюється через бібліотеку requests.

Для керування конфігураціями і секретами використовується `dotenv`, а логування подій у системі організоване за допомогою стандартного модуля `logging`. Тестування системи проводиться через окремі тестові скрипти, які перевіряють коректність роботи API та модулів обробки даних. Код організований за модульною структурою з окремими папками для моделей, сервісів, утиліт, шаблонів і статичних файлів, що забезпечує зручність підтримки та розвитку проєкту. Такий стек і підхід відповідають сучасним стандартам розробки на Python і дозволяють ефективно працювати з екологічними мультиспектральними даними. Модуль збору та обробки даних у цьому проєкті відповідає за отримання, аналіз. Основні функції модуля включають збір даних, обробку, аналіз, візуалізацію та повернення результатів.

Збір даних здійснюється через отримання мультиспектральних супутникових даних за допомогою зовнішніх API, таких як Copernicus/Sentinel, а також завантаження NetCDF-файлів та інших форматів із відкритих джерел. Для обробки даних використовується бібліотека `xarray`, що дозволяє відкривати та читати багатовимірні наукові дані. Вибір пріоритетних змінних для аналізу включає основні атмосферні компоненти – NO_2 , O_3 , CO , SO_2 , HCHO , CH_4 , аерозолі тощо. Дані проходять фільтрацію, очищення та підготовку до подальшого аналізу.

Аналіз передбачає розрахунок базових статистичних показників: середнього, медіани, мінімуму, максимуму, стандартного відхилення та квантилів. Для візуалізації будуються просторові карти (`heatmap`) та гістограми розподілу значень, а також застосовується оцінка густини розподілу (`kernel density estimation`) для глибшого аналізу.

Результати формуються у вигляді структурованих словників (`dict`), які використовуються у веб-інтерфейсі або API. Для відображення на фронтенді генеруються візуалізації у форматі зображень (`base64`). Такий підхід забезпечує комплексну обробку екологічних даних і їх ефективне представлення користувачам.

Висновки до третього розділу

Описано реалізовану інтеграцію з Copernicus API та іншими зовнішніми сервісами для автоматизованого збору мультиспектральних супутникових і метеорологічних даних, створено клієнт для формування запитів до STAC API, модуль збору, обробки та аналізу екологічної інформації з використанням сучасного Python-стеку, а також забезпечено візуалізацію результатів і можливість оперативного сповіщення користувачів про екологічні загрози.

ВИСНОВКИ

Розроблена система автоматизує процес збору, обробки та аналізу екологічних даних із супутникових і метеорологічних джерел, що дозволяє оперативно отримувати інформацію про стан атмосфери та рівень забруднення повітря. Актуальність такої розробки обумовлена зростанням екологічних викликів, необхідністю моніторингу якості повітря у містах та регіонах, а також потребою у сучасних цифрових інструментах для прийняття рішень у сфері охорони довкілля.

Система має модульну архітектуру, легко масштабується, підтримує інтеграцію з різними зовнішніми сервісами та може бути розширена для впровадження алгоритмів машинного навчання, прогнозування та автоматичного сповіщення користувачів.

Перспективи впровадження:

- інтеграція з додатковими джерелами даних (NASA, локальні метеостанції);
- додавання модулів машинного навчання для прогнозування та виявлення аномалій;
- використання у міських системах моніторингу, наукових дослідженнях, екологічних стартапах;
- розширення функціоналу для автоматичного сповіщення органів влади та населення про критичні ситуації;
- система є сучасним інструментом для екологічного моніторингу, що має значний потенціал для подальшого розвитку та впровадження у різних сферах.

СПИСОК ВИКОРИСТАНИХ ДЖЕРЕЛ

1. Ambient (outdoor) air pollution. World Health Organization (WHO). URL: [https://www.who.int/news-room/fact-sheets/detail/ambient-\(outdoor\)-air-quality-and-health](https://www.who.int/news-room/fact-sheets/detail/ambient-(outdoor)-air-quality-and-health) (date of access: 24.03.2025).
2. Roser M. Data review: how many people die from air pollution?. Our World in Data. URL: <https://ourworldindata.org/data-review-air-pollution-deaths> (date of access: 24.03.2025).
3. Екологічні проблеми промислового комплексу - Бібліотека BukLib.net. Головна - Бібліотека BukLib.net. URL: <https://buklib.net/books/23380/> (дата звернення: 24.03.2025).
4. «Створення ефективної системи моніторингу довкілля в Україні: проблеми і шляхи їх вирішення». Аналітична записка. Національний інститут стратегічних досліджень. URL: <https://niss.gov.ua/doslidzhennya/nacionalna-bezpeka/stvorenniya-efektivnoi-sistemi-monitoringu-dovkillya-v-ukraini> (дата звернення: 31.03.2025).
5. 10 найбрудніших підприємств України: половина розташована на Донеччині - Хмарочос. Хмарочос. URL: <https://hmarochos.kiev.ua/2020/08/10/10-najbrudnishyh-pidpryyemstv-ukrayiny-polovyna-roztashovana-na-donechchyni/> (дата звернення: 01.04.2025).
6. Нова державна система моніторингу довкілля збиратиме та узагальнюватиме екологічні дані з усієї України – Міністерство захисту довкілля та природних ресурсів України. Міністерство захисту довкілля та природних ресурсів України – офіційний сайт. URL: <https://mepr.gov.ua/nova-derzhavna-systema-monitoryngu-dovkillya-zbyratyme-ta-uzagalnyuvatyme-ekologichni-dani-z-usiyei-ukrayiny/> (дата звернення: 01.04.2025).

7. Супутникові знімки: принципи, застосування та майбутні тенденції. TS2 Space. URL: <https://ts2.tech/uk/супутникові-знімки-принципи-застосу/> (дата звернення: 04.04.2025).
8. Sentinel-5P - Atmospheric measurements with high spatio-temporal resolution. Open-source satellietbeelden? Slim bekeken!. URL: <https://terrascop.be/en/satellites/sentinel-5p> (date of access: 04.04.2025).
9. Моніторинг якості атмосферного повітря над територією України з деталізацією для міст за даними супутника Sentinel-5P | Visnik Nacional noi akademii nauk Ukraini. Visnik Nacional noi akademii nauk Ukraini. URL: <https://visnyk-nanu.org.ua/ojs/index.php/v/article/view/1> (дата звернення: 04.04.2025).
- 10.10 найкращих інструментів штучного інтелекту для моніторингу навколишнього середовища (червень 2025 р.). Unite.AI. URL: <https://www.unite.ai/uk/найкращі-інструменти-штучного-інтелекту-для-моніторингу-навколишнього-середовища/> (дата звернення: 04.04.2025).
11. Головна. URL: <http://ecoj.dea.kiev.ua/archives/2023/1/28.pdf> (дата звернення: 04.04.2025).
12. Особливості використання технологій штучного інтелекту в управлінні природними ресурсами та забезпеченні екологічної безпеки | Науковий вісник Ужгородського національного університету. Серія: Право. Науковий вісник Ужгородського національного університету. Серія: Право. URL: <http://visnyk-pravo.uzhnu.edu.ua/article/view/329995> (дата звернення: 04.04.2025).
13. Штучний інтелект та зміна клімату: прогнозування екологічних тенденцій. The Future of AI Companies: Human-AI Collaboration at Neuron.expert. URL: <https://neuron.expert/news/ai-and-climate-change-predicting-environmental-trends/13521/uk/> (дата звернення: 04.04.2025).
14. Якість повітря в реальному часі та прогноз шкідливих викидів з заводів та пожеж - YourAirTest. Your Air Test. URL: <https://yourairtest.com/uk/> (дата звернення: 04.04.2025).

15. Аналіз геопросторових даних. DSpace :: ELAKPI :: Репозитарій КПІ ім. Ігоря Сікорського. URL: <https://ela.kpi.ua/server/api/core/bitstreams/213801fb-45e7-4fd2-80ac-ed2b17c41e3a/content> (дата звернення: 04.04.2025).
16. Проф., д-р Петро Когут. Просторовий Аналіз: Обробка ГІС Даних та Їх Використання. EOS Data Analytics. URL: <https://eos.com/uk/blog/prostorovyi-analiz/> (дата звернення: 04.04.2025).
17. Національний фонд досліджень України – Віримо в нашу перемогу!. URL: https://nrfu.org.ua/wp-content/uploads/2022/01/2020.01_0268_kussul_083_01.2020_zz-1-stysnuto.pdf (дата звернення: 04.04.2025).
18. Головна | Elib LNTU. URL: https://elib.lntu.edu.ua/sites/default/files/elib_upload/Кондіус%202%20готовва/page15.html (дата звернення: 04.02.2025).
19. Діаграми UML для моделювання процесів і архітектури проекту. Evergreen - web розробка і діджиталізація бізнесу за допомогою AI продуктів. URL: <https://evergreens.com.ua/ua/articles/uml-diagrams.html> (дата звернення: 02.02.2025).
20. Welcome To UML Web Site!. Welcome To UML Web Site!. URL: <https://www.uml.org/> (date of access: 14.02.2025).
21. Unified Modeling Language (UML) description, UML diagram examples, tutorials and reference for all types of UML diagrams - use case diagrams, class, package, component, composite structure diagrams, deployments, activities, interactions, profiles, etc. Unified Modeling Language (UML) description, UML diagram examples, tutorials and reference for all types of UML diagrams - use case diagrams, class, package, component, composite structure diagrams, deployments, activities, interactions, profiles, etc. URL: <https://www.uml-diagrams.org/> (date of access: 14.02.2025).
22. Homepage | Copernicus. Homepage | Copernicus. URL: <https://www.copernicus.eu/en> (date of access: 14.01.2025).

- 23.APIs. Copernicus Data Space Ecosystem. URL: <https://dataspace.copernicus.eu/analyse/apis> (date of access: 14.01.2025).
- 24.GitHub - armkhudinyan/copernicus_api: Copernicus API makes searching and downloading of Copernicus Sentinel mission images from Copernicus Data Space Ecosystem (CDSE) easy. GitHub. URL: https://github.com/armkhudinyan/copernicus_api (date of access: 07.01.2025).
- 25.Accessing Sentinel mission data via the new Copernicus Data Space Ecosystem APIs. Copernicus Data Space Ecosystem. URL: <https://dataspace.copernicus.eu/news/2023-9-28-accessing-sentinel-mission-data-new-copernicus-data-space-ecosystem-apis> (date of access: 11.01.2025).
- 26.How to open and visualize Copernicus Marine Data using Python? | Copernicus Marine Help Center. Copernicus Marine Help Center. URL: <https://help.marine.copernicus.eu/en/articles/4854800-how-to-open-and-visualize-copernicus-marine-data-using-python> (date of access: 21.01.2025).
- 27.API Python: Copernicus Data Hub. STEP Forum. URL: <https://forum.step.esa.int/t/api-python-copernicus-data-hub/25694> (date of access: 17.01.2025).
- 28.Climate Data Store. Climate Data Store. URL: <https://cds.climate.copernicus.eu/how-to-api> (date of access: 02.01.2025).
- 29.Examples Documentation. Copernicus Data Space Ecosystem. URL: <https://documentation.dataspace.copernicus.eu/APIs/SentinelHub/BatchV2/Examples.html> (date of access: 01.01.2025).
- 30.New Tutorial Videos for Sentinel Hub API Newcomers. Copernicus Data Space Ecosystem. URL: <https://dataspace.copernicus.eu/news/2024-1-24-new-tutorial-videos-sentinel-hub-api-newcomers> (date of access: 05.01.2025).
- 31.SQLAlchemy. SQLAlchemy - The Database Toolkit for Python. URL: <https://www.sqlalchemy.org/> (date of access: 12.02.2025).

32. Національний університет біоресурсів і природокористування України.
URL: <https://nubip.edu.ua/sites/default/files/u243/24.pdf> (дата звернення: 09.04.2025).
33. DSpace :: ELAKPI :: Репозитарій КПІ ім. Ігоря Сікорського. URL:
<https://ela.kpi.ua/server/api/core/bitstreams/6b5778a5-942c-46f5-a148-625f87753a7a/content> (дата звернення: 17.04.2025).
34. Проф., д-р Петро Когут. Просторовий Аналіз: Обробка ГІС Даних та Їх Використання. EOS Data Analytics. URL:
<https://eos.com/uk/blog/prostorovyi-analiz/> (дата звернення: 14.01.2025).
35. Проф., д-р Петро Когут. Просторовий Аналіз: Обробка ГІС Даних та Їх Використання. EOS Data Analytics. URL:
<https://eos.com/uk/blog/prostorovyi-analiz/> (дата звернення: 30.03.2025).

ДОДАТКИ

ДОДАТОК А

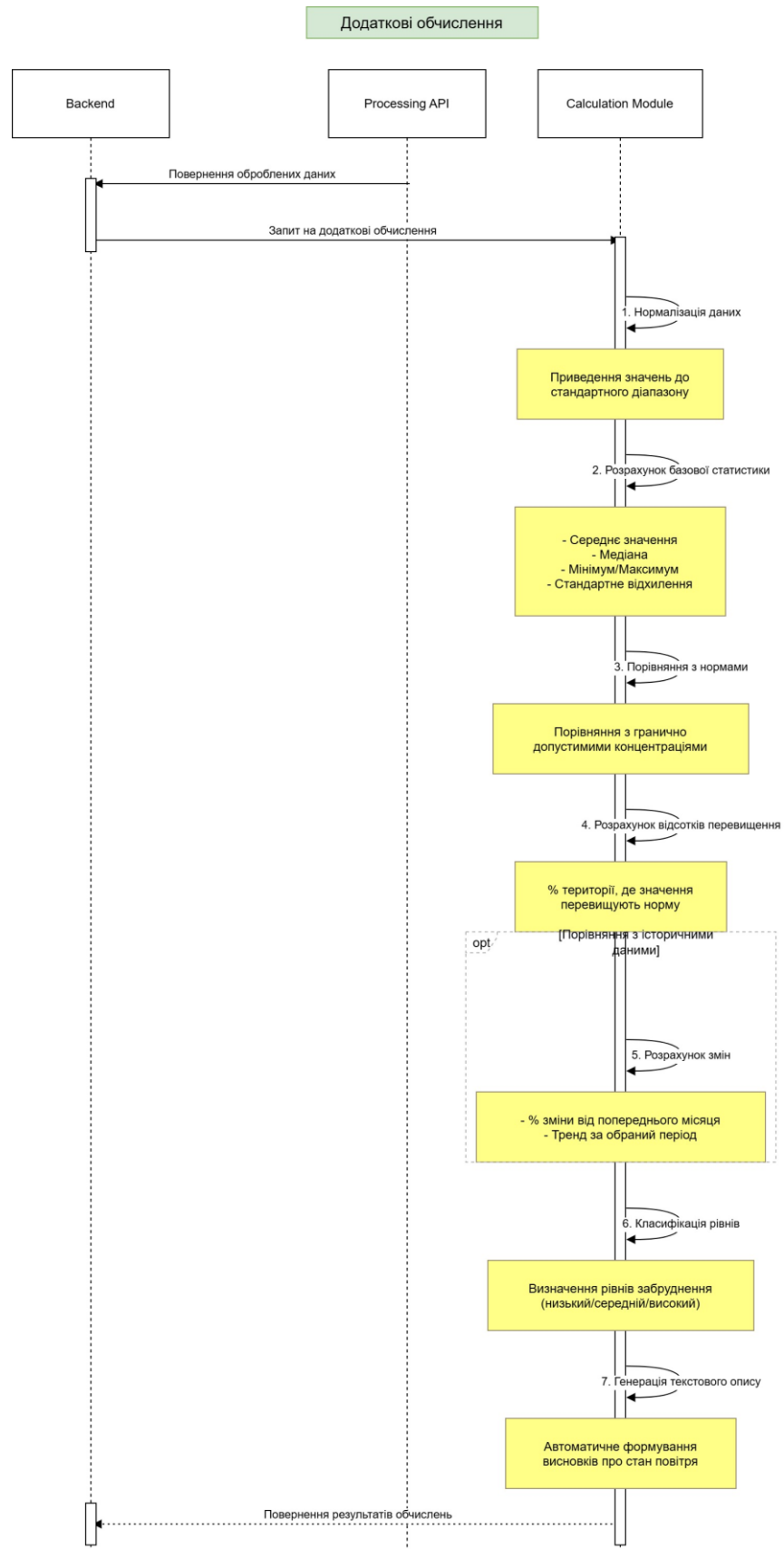


Рисунок А1 – Продовження UML-діаграми послідовностей