

**МІНІСТЕРСТВО ОСВІТИ І НАУКИ УКРАЇНИ
НАЦІОНАЛЬНИЙ ТЕХНІЧНИЙ УНІВЕРСИТЕТ УКРАЇНИ
“КИЇВСЬКИЙ ПОЛІТЕХНІЧНИЙ ІНСТИТУТ
імені Ігоря Сікорського”
НАВЧАЛЬНО-НАУКОВИЙ ФІЗИКО-ТЕХНІЧНИЙ ІНСТИТУТ
Кафедра математичного моделювання та аналізу даних**

До захисту допущено:

Завідувач кафедри

Куcssуль Н.М.

«___»___2022 р.

Дипломна робота

на здобуття ступеня бакалавра
за освітньо-професійною програмою «Математичні
методи моделювання, розпізнавання образів та
безпеки даних»

спеціальності: 113 «Прикладна математика»

на тему: «Хмарні технології для класифікації земного покриття та оцінки
індикаторів сталого розвитку в Україні»

Виконав: здобувач вищої освіти IV курсу, групи ФІ-82 Пилип Лідія
Миколаївна

Керівник: професор, доктор технічних наук Шелестов А.Ю.

Рецензент: с.н.с., к.т.н., Яйлимов Б.Я. _____

Засвідчую, що у цій дипломній роботі
немає запозичень із праць інших
авторів без відповідних посилань.

Студент _____

НАЦІОНАЛЬНИЙ ТЕХНІЧНИЙ УНІВЕРСИТЕТ УКРАЇНИ
«КИЇВСЬКИЙ ПОЛІТЕХНІЧНИЙ ІНСТИТУТ
імені ІГОРЯ СІКОРСЬКОГО»
ФІЗИКО-ТЕХНІЧНИЙ ІНСТИТУТ

Кафедра математичного моделювання та аналізу даних

Рівень вищої освіти — перший (бакалаврський)

Спеціальність (освітня програма) — 113 Прикладна математика,

ОПП «Математичні методи моделювання, розпізнавання образів та безпеки даних»

ЗАТВЕРДЖУЮ

Завідувач кафедри

Наталія КУССУЛЬ

«__» _____ 2022 р.

ЗАВДАННЯ

на дипломну роботу

Студент: Пилип Лідія Миколаївна

1. Тема роботи: «Хмарні технології для класифікації земного покриття та оцінки індикаторів сталого розвитку в Україні»

Керівник:

Професор, Доктор технічних наук Шелестов Андрій Юрійович,

затверджені наказом по університету №__ від «__» _____ 2022 р.

2. Термін подання студентом роботи: «__» _____ 2022 р.
3. Вихідні дані до роботи: методи класифікації земного покриття та оцінювання індикаторів цілей сталого розвитку
4. Зміст роботи: аналіз існуючих задач та підходів та джерел даних; методи розв'язання задач класифікації та оцінювання індикаторів SDG; поточний стан розв'язання прикладних задач в області геопросторового аналізу.
5. Перелік ілюстративного матеріалу (із зазначенням плакатів, презентацій тощо): Презентація доповіді
6. Дата видачі завдання: 10 вересня 2021 р.

Календарний план

№ з/п	Назва етапів виконання дипломної роботи	Термін виконання	Примітка
1	Узгодження теми роботи з науковим керівником	20-30 жовтня 2021 р.	Виконано
2	Огляд опублікованих джерел за тематикою дослідження	листопад-грудень 2021 р	Виконано
3	Написання 1-го розділу	Січень 2022 р.	Виконано
4	Написання 2-го розділу	Лютий 2022 р.	Виконано
5	Написання 3-го розділу	Березень 2022 р.	Виконано
6	Написання 4-го розділу	Квітень 2022 р	Виконано

7	Написання вступу, висновків та анотацій	Травень 2022 р.	Виконано
8	Здача готової роботи на кафедру	13.06.2022 р.	Виконано
9	Попередній захист на кафедрі	14.06.2022 р.	Виконано
10	Подання завершеної роботи до Державної екзаменаційної комісії		

Студент

Пилип Л.М.

Керівник

Шелестов А.Ю.

РЕФЕРАТ

Кваліфікаційна робота містить: 55 стор., 16 рисунків, 4 таблиці, 43 джерела.

Сільськогосподарська галузь України є одним з основних секторів зростання української економіки. Тому дуже важливо проводити якісний аналіз земного покриття, а також здійснювати оцінювання індикаторів цілей сталого розвитку (Sustainable Development Goal — SDG), на основі чого можна буде приймати зважені стратегічні рішення. На цьому тлі національні експерти зіткнулися з проблемою великих даних у сфері спостереження Землі. На щастя, хмарні рішення, такі як Amazon Web Services, Google Earth Engine та інші платформи, забезпечують доступ до даних Sentinel-1, Sentinel-2 і Landsat з просторовою роздільною здатністю від 10 до 30 м, можливістю швидкого та зручного способу обробки геопросторових даних та їх ефективне використання для побудови різноманітних інформаційних продуктів, таких як карти класифікації ґрунтового покриття, моніторинг стану посівів тощо.

За останні декілька років українські експерти мали досвід використання хмарних технологій у наукових та інноваційних проектах, які підтримуються Європейською Комісією, Світовим Банком, Програмою розвитку ООН, Комітетом GEO, а також досвід розробки програмного забезпечення з відкритим кодом, зокрема на основі програмного забезпечення Open Data Cube, та використання методів машинного навчання у хмарному середовищі.

Розроблені технології реалізуються на різноманітних хмарних платформах і дозволяють розв'язати різного роду прикладні задачі, зокрема пов'язані з моніторингом сільськогосподарських угідь, оцінкою індикаторів цілей сталого

розвитку на національному рівні. Усі ці питання будуть більш детально описані в цій роботі.

ХМАРНІ ТЕХНОЛОГІЇ, МОНІТОРИНГ СІЛЬСЬКОГО ГОСПОДАРСТВА,
СУПУТНИКОВИЙ МОНІТОРИНГ, ЦІЛІ СТАЛОГО РОЗВИТКУ

ABSTRACT

Qualification work contains: 55 pages, 16 figures, 4 tables, 43 sources.

Ukraine's agricultural sector is one of the main sectors of growth of the Ukrainian economy. Therefore, it is very important to conduct a qualitative analysis of the land cover, as well as to evaluate the indicators of sustainable development goals (SDG), on the basis of which it will be possible to make informed strategic decisions. Against this background, national experts faced the problem of big data in the field of Earth observation. Fortunately, cloud solutions such as Amazon Web Services, Google Earth Engine, and other platforms provide access to Sentinel-1, Sentinel-2, and Landsat data with a spatial resolution of 10 to 30 m, and a fast and convenient way to process geospatial data. and their effective use for the construction of various information products, such as soil classification maps, crop condition monitoring, etc.

Over the past few years, Ukrainian experts have experienced the use of cloud technologies in research and innovation projects supported by the European Commission, the World Bank, the United Nations Development Program, the GEO Committee, as well as experience in developing open source software, including Open Source Data. , and the use of machine learning methods in a cloud environment.

The developed technologies are implemented on various cloud platforms and tasks to solve various applications, special kinds, related to the monitoring of farmland, evaluation of indicators of sustainable development goals at the national level. All these issues will be described in more detail in this paper.

CLOUD TECHNOLOGIES, AGRICULTURAL MONITORING, SATELLITE
MONITORING, SUSTAINABLE DEVELOPMENT GOALS

ЗМІСТ

ВСТУП.....	10
РОЗДІЛ 1 ОГЛЯД ІСНУЮЧИХ ЗАДАЧ ТА ПІДХОДІВ.....	13
1.1 Проекти Світового банку.....	13
1.2 Проект Amazon-GEO «Методологія оцінки індикаторів SDG» ..	15
1.3 Проект GEO-GEE «Супутниковий моніторинг для сталого землеустрою та сільського господарства в Україні за допомогою Google Earth Engine».	19
Висновки до розділу 1	21
РОЗДІЛ 2 ДЖЕРЕЛА ДАНИХ.....	22
2.1. Супутникові дані	22
2.2 Наземні дані	25
Висновки до розділу 2.....	27
РОЗДІЛ 3 ЗАДАЧА КЛАСИФІКАЦІЇ ТА ОЦІНКА ІНДИКАТОРІВ ЦІЛЕЙ СТАЛОГО РОЗВИТКУ	28
3.1 Класифікація земного покриття	28
3.2 Оцінка індикаторів цілей сталого розвитку.....	31
3.2.1 Оцінювання індикатора SDG 15.1.1 “Forest area as proportion of total land area”.....	33
3.2.2 Оцінювання індикатора SDG 15.3.1	34
3.2.3 Оцінювання індикатора SDG 11.3.1 “Ratio of land consumption rate to population growth rate”	36

3.2.4 Оцінювання індикатора SDG 2.4.1 “Proportion of agricultural area under productive and sustainable agriculture”	37
Висновки до розділу 3	38
РОЗДІЛ 4 РЕЗУЛЬТАТИ РОЗВ’ЯЗАНИХ ПРИКЛАДНИХ ЗАДАЧ.....	39
4.1 Результати побудови карт класифікації	39
4.2 Результати оцінки індикаторів SDG	45
Висновки до розділу 4	47
ВИСНОВКИ	48
ПЕРЕЛІК ВИКОРИСТАНИХ ДЖЕРЕЛ	50

ВСТУП

Актуальність дослідження. За останні п'ять років сільськогосподарське виробництво в Україні стало одним з основних секторів зростання української економіки. Сьогодні переважна більшість експортних доходів породжена саме сільським господарством. Ця галузь також сприяє розвитку пов'язаних та взаємодоповнюючих галузей промисловості, дозволяє Україні займати гідне місце на міжнародному ринку. Тим не менше, рівень діджиталізації України є значно нижчим, ніж у країнах Європейського Союзу. Очевидно, що з розвитком цифрових технологій, застосуванням їх для прийняття бізнес-рішень, розробки стратегій збільшується й прибутковість справи, підвищується рівень довіри клієнтів. Також українській сільськогосподарській політиці бракує складних інструментів управління, таких як Land Parcel Identification System (LPIS), яка в Європі працює вже з 2013 року.

Іншою важливою сферою діяльності, якою зараз займаються всі розвинені країни, є оцінювання індикаторів цілей сталого розвитку. Так, у постанові зі сталого розвитку на період до 2030 року, прийнятій усіма членами ООН у 2015 році, акцент робиться на SDG та відповідних індикаторах, що характеризуються складною взаємодією людини та навколишнього середовища. Чимало показників базуються на використанні геопросторової інформації та можуть бути отримані із супутникових даних, що інтегровані разом із вимірюваннями та моделями, включно з моделями метеорологічного прогнозування, біофізичними параметрами рослинності та моделями класифікації (машинного навчання). Тому виникає потреба у розробці методів та інструментів, які дозволять отримувати оцінки індикаторів SDG у вигляді геоінформаційних продуктів.

Метою дослідження є аналіз підходів до діджиталізації процесів моніторингу земельних ресурсів України за допомогою хмарних технологій, а також методів оцінювання індикаторів цілей сталого розвитку, які стосуються земного покриття. Для досягнення мети необхідно дослідити наявні результати досліджень на дану тематику та проаналізувати кроки подальшого розвитку. Для вирішення цієї задачі потрібно виконати такі завдання:

- 1) провести аналіз опублікованих результатів досліджень, пов'язаних з розв'язанням задачі класифікації земного покриття на основі хмарних технологій, а також методів оцінювання індикаторів цілей сталого розвитку;
- 2) дослідити наявні джерела даних, які можуть бути використані для таких досліджень та роль використання хмарних технологій для їх опрацювання;
- 3) розглянути наявну методологію для здійснення класифікації земного покриття й оцінки індикаторів SDG;
- 4) проаналізувати та порівняти результати використання різних моделей, методів та технологій.

Об'єктом дослідження є методи класифікації земного покриття та оцінювання індикаторів SDG.

Предметом дослідження є геопросторові дані з різних джерел.

При розв'язанні поставлених завдань використовувались такі методи дослідження, як-от: системний аналіз, інтелектуальний аналіз даних.

Наукова новизна отриманих результатів полягає в тому, що вперше було здійснено дослідження, яке містить в собі результати аналізу методів інтелектуальної обробки даних, зокрема нейронних мереж, методів оцінки індикаторів цілей сталого розвитку та їх реалізації на базі хмарних технологій у

різноманітних глобальних та національних проектах, націлених на класифікацію земного покриву та оцінювання індикаторів SDG.

Практичне значення результатів полягає у формуванні підґрунтя для ефективного використання сучасних методів обробки геопросторових даних різної природи із застосуванням хмарних технологій, що в майбутньому може бути використано для подальшого розвитку прикладної області інтелектуальних обчислень.

РОЗДІЛ 1 ОГЛЯД ІСНУЮЧИХ ЗАДАЧ ТА ПІДХОДІВ

Експертна група, що складається із спеціалістів кафедри математичного моделювання та аналізу даних Науково-навчального Фізико-технічного інституту КПІ імені Ігоря Сікорського, а також Інституту космічних досліджень НАН та ДКА України, є учасником програм Joint Experiment for Crop Assessment and Monitoring (JECAM) та Group on Earth Observations' Global Agricultural Monitoring Initiative (GEOGLAM). Експерти беруть участь у різноманітних міжнародних та національних проектах, а також проводять моніторингові дослідження більше 10 років в інтересах держави та ряду міжнародних організацій, включаючи комітет зі спостереженні Землі (Group on Earth Observations — GEO).

Основною метою всіх цих програм та проектів є позитивний вплив на економіку та надання нових можливостей прийняття управлінських рішень на рівні держави на основі результатів супутникового моніторингу.

1.1 Проекти Світового банку

Проекти Світового банку підтримують широкий спектр інвестицій у таких галузях, як освіта, охорона здоров'я, державне управління, інфраструктура, розвиток фінансового та приватного секторів, сільське господарство, а також природокористування та управління природними ресурсами [1].

У липні 2021 року в Україні відкрився ринок землі [2]. Для забезпечення прозорості, справедливості та надійності цього процесу необхідна об'єктивна інформація про реальне землекористування та його історію, стан розвитку посівів

у кожен вегетаційний період. Усі ці показники істотно впливають на ринкову вартість землі.

Програма Світового банку за фінансування Європейським Союзом та іншими міжнародними партнерами в межах програми «Підтримка прозорого управління земельними ресурсами в Україні» сприяє впровадженню земельної реформи в Україні. Тривалість програми становить 5 років та передбачає допомогу у врегулюванні земельних відносин, забезпеченні повноти та коректності відображення актуальної інформації щодо наявних земельних ресурсів та забезпеченні вільного доступу до цієї інформації, покращенні захисту прав землевласників та землекористувачів, а також аналітичної підтримки політичного діалогу щодо земельної реформи.

У межах проектів «Підтримка прозорого управління земельними ресурсами в Україні» (2016–2019, 2021) та «Обробка супутникових зображень для аналізу земного покриття», які реалізуються в межах проектів «Support to Agriculture and Food Policy Implementation (SAFPI)» та «Moving Forward Together» (2020), що фінансуються ЄС, було проаналізовано доцільність використання хмарної технології Google Earth Engine для виконання таких задач:

- 1) моніторингу використання землі;
- 2) можливості розробки та використання рішення для аналізу стану посівів;
- 3) розробки LC/LU мап, що демонструють різноманітні фізичні та цільові характеристики територій, надають цінну геоінформацію для просторового розподілу різних земельних класів та їх тимчасових змін.

Цей аналіз виконується з використанням вільних інноваційних хмарних технологій, відкритих джерел супутникової інформації (Sentinel-1 і Sentinel-2), а також моделей штучного інтелекту, які використовуються для побудови карт

класифікації земного покриву. Для даного проекту на платформі Google Earth Engine використано метод Random Forest Classifier (RF) — один з найпопулярніших методів машинного навчання, що полягає у застосуванні ансамблю дерев рішень, котрий використовується для вирішення задач класифікації та регресії.

Зокрема, в ході роботи над проектом автори запропонували використати раніше розроблений підхід для створення на рівні держави LC/LU карт високого просторового розрізнення, які представляють різні фізичні та цільові характеристики міських територій, надають цінну геоінформацію для просторового розподілу різних класів та їх тимчасових змін. Запропонований підхід перевершує глобальні продукти в питаннях точності та просторового розрізнення.

1.2 Проект Amazon-GEO «Методологія оцінки індикаторів SDG»

Разом з іншими країнами-членами ООН Україна приймає участь у глобальному процесі забезпечення сталого розвитку. Керуючись принципом «нікого не залишити осторонь», на період до 2030 року було розпочато процес адаптації цілей сталого розвитку, який вимагає глибоких соціально-економічних змін в Україні та нового глобального партнерства [3].

На сьогоднішній день методики розрахунку показників SDG в основному базуються на супутникових даних з грубим просторовим розрізненням (300 м) та глобальних продуктах. Однак попередні дослідження для України показали, що глобальні продукти з таким просторовим розрізненням не є дуже точними, особливо коли йдеться про локальний масштаб (на рівні країни). Особливо це стосується оцінювання потенціалу сільського господарства через створення

сумішей пікселів з комбінування різних типів ґрунтового покриття в одному пікселі [6]. Саме тому в рамках спільного гранту компанії Amazon та комітету GEO «Методика оцінки показників SDG» автори розробили автоматизовану технологію розрахунку індикаторів SDG на хмарній платформі Amazon Web Services з використанням сучасних технологій Open Data Cube [7] (Рис. 1.1).

Для оцінки сільськогосподарських площ було запропоновано поєднати цей підхід із біофізичною моделлю WOFOST (WOrld FOod STudies) [8] як альтернативним джерелом інформації, що дозволяє підвищити просторове розрізнення показників стану посівів.

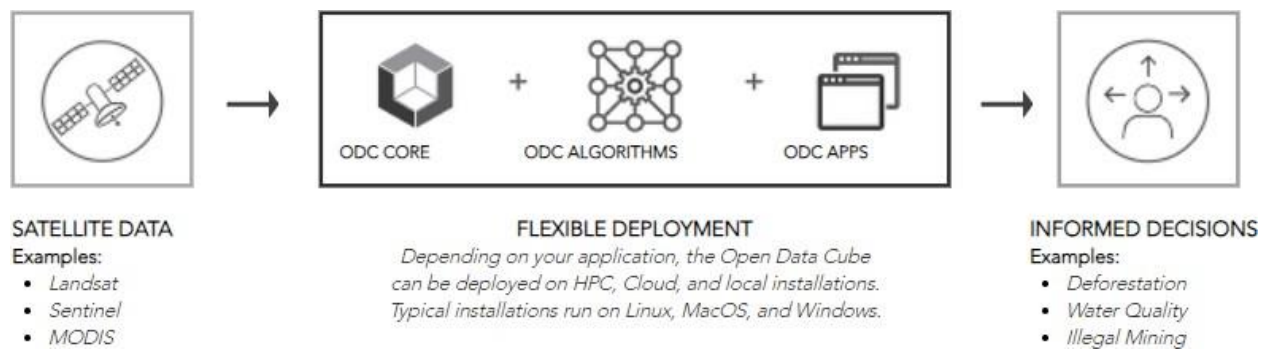


Рисунок 1.1 – Екосистема ODC

Працюючи над проектом Amazon-GEO «Методика оцінки показників SDG» (2019–2022 рр.), українська команда розробила автоматизовані робочі процеси для розрахунку індикаторів SDG, згрупованих за класами ґрунтового покриття. Основна мета – прийняти, удосконалити та застосувати вже запропоновані методології, які використовувалися для генерування глобальних продуктів від даних низького просторового розрізнення до даних з більшим просторовим розрізненням (10 м), які краще підходять для регіональних продуктів і додатків.

У рамках проекту були досліджені такі індикатори SDG:

- 1) 2.4.1 «Частка сільськогосподарських площ під продуктивним сталим сільським господарством» [9] — надає оцінку прогресу в напрямку сталого сільського господарства;
- 2) 11.3.1 «Співвідношення рівня використання землі до темпів зростання населення» [10] — вимірює, наскільки ефективно міста використовують землю, що вимірюється як відношення швидкості, з якою міста просторово використовують землю, до швидкості зростання їх населення;
- 3) 15.1.1 «Площа лісів як частка загальної площі земель» [11] — виражається у відсотках;
- 4) 15.3.1 «Частка деградованих земель серед загальної площі землі» [12] — дозволяє оцінити масштаби деградації земель та, як і індикатор 15.3.1, необхідний в рамках завдання боротьби з опустелюванням, відновлювання деградованих земель та ґрунтів, включаючи землі, постраждалі від опустелювання, посухи та повеней.

Інформаційна технологія для цих індикаторів реалізована в хмарному середовищі Amazon Web Services (AWS) і була заснована на технології Open Data Cube (проект програмного забезпечення для управління та аналізу геопросторових даних з відкритим вихідним кодом, який спрямований на допомогу у використанні можливостей супутникових даних), а також алгоритму глибокого навчання для класифікації LC/LU, біофізичного моделювання, метеорологічного моделювання та аналізу супутникових даних. Ця технологія є масштабованою і придатною для використання в будь-якій іншій країні або регіоні.

Основна інновація проекту пов'язана з удосконаленням існуючих робочих процесів для розрахунку індикаторів SDG шляхом використання даних високого просторового розрізнення та заповнення прогалів між існуючими глобальними та

національними продуктами. Зокрема, в роботі «Deep recurrent neural network for crop classification task based on Sentinel-1 and Sentinel-2 imagery» [13] запропоновано використовувати раніше розроблену неймережеву модель для побудови карти класифікації LC/LU високого просторового розрізнення на рівні країни. Приклад карти класифікації, отриманої в межах даного проекту, представлено на Рис. 1.2.

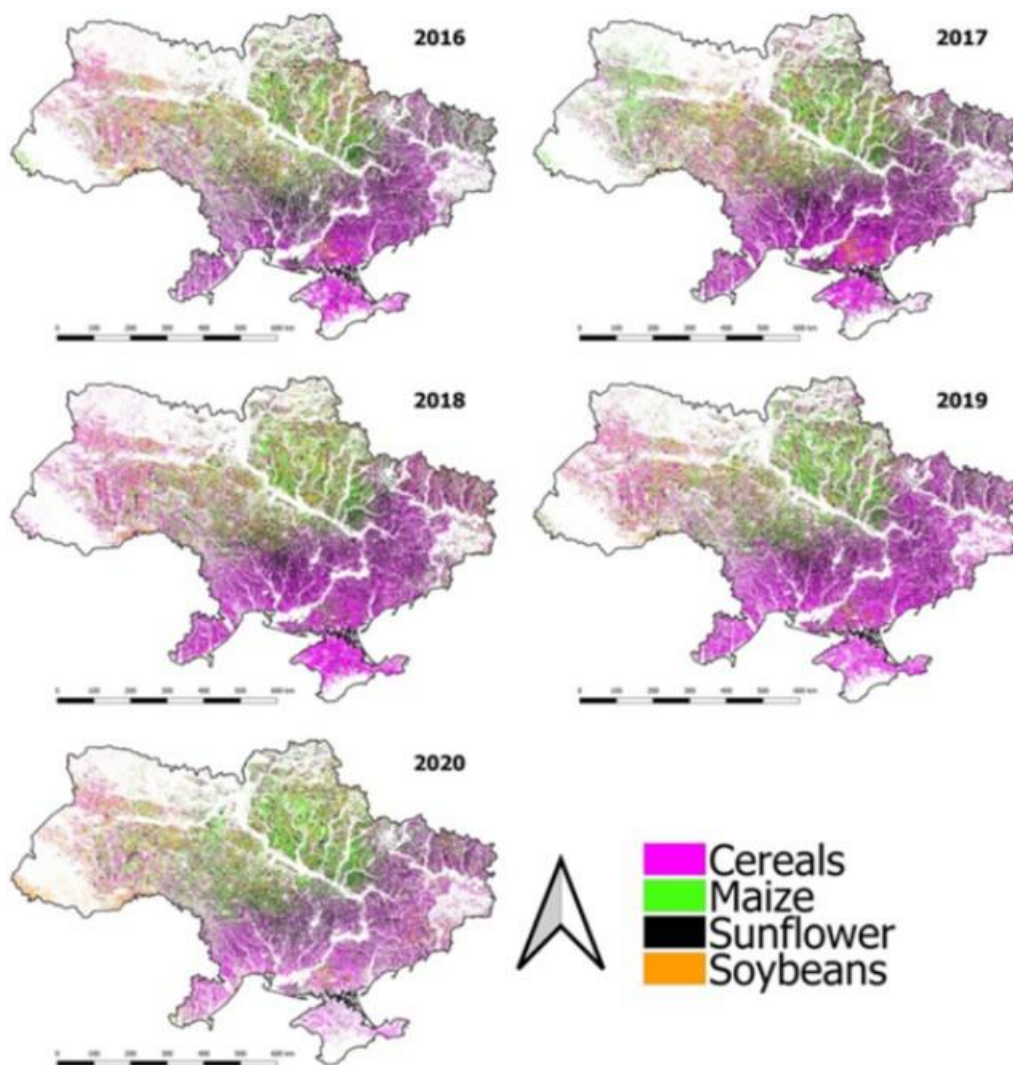


Рисунок 1.2 – Класифікаційні мапи найпоширеніших культур в Україні, 2016-2020 рр.

Запропонований підхід перевершує глобальні продукти за точністю та просторовим розрізненням. Для оцінки площі сільськогосподарських культур в умовах продуктивного та стійкого сільського господарства пропонується поєднати цей підхід з біофізичною моделлю WOFOST як альтернативним джерелом інформації.

1.3 Проект GEO-GEE «Супутниковий моніторинг для сталого землеустрою та сільського господарства в Україні за допомогою Google Earth Engine»

Проект «Супутниковий моніторинг для сталого землеустрою та сільського господарства в Україні з використанням Google Earth Engine» (2020-2022 рр.) був номінований спільною програмою Google Earth Engine та комітетом GEO. В межах проекту учасники консорціуму отримали можливість ліцензованого доступу до платформи Google Cloud для проведення наукових досліджень, що надає набір загальнодоступних хмарних послуг, які пропонує компанія Google, та включає в себе цілий ряд хмарних служб для обчислень, зберігання та розробки додатків. Відкрита платформа Google Earth Engine надає доступ до традиційних методів машинного навчання та попиксельної класифікації.

Сьогодні методи глибокого навчання перевершують традиційні методи машинного навчання при розв'язанні багатьох прикладних задач, таких як переклад людської мови, задачі комп'ютерного зору, розпізнавання мовлення тощо. Для аналізу даних дистанційного зондування Землі методи глибокого навчання також застосовуються (Рис. 1.3), але тут починають з'являтися труднощі та проблеми. Однією з часткових проблем є своєчасна класифікація ділянок, яка вимагає широкого використання часових рядів багатосенсорних просторових даних. Крім

того, це завдання критично залежить від інтеграції геопросторової інформації (наприклад, сегментації ділянок) та асиміляції додаткових даних із зовнішніх джерел (наприклад, метеорологічна інформація, агрономічні знання на практиці). Наразі основні дані для задач навчання та валідації результатів, як правило, є недостатніми та не мають єдиної універсальної форми, особливо на рівні великих виробничих зон або на рівні країни.

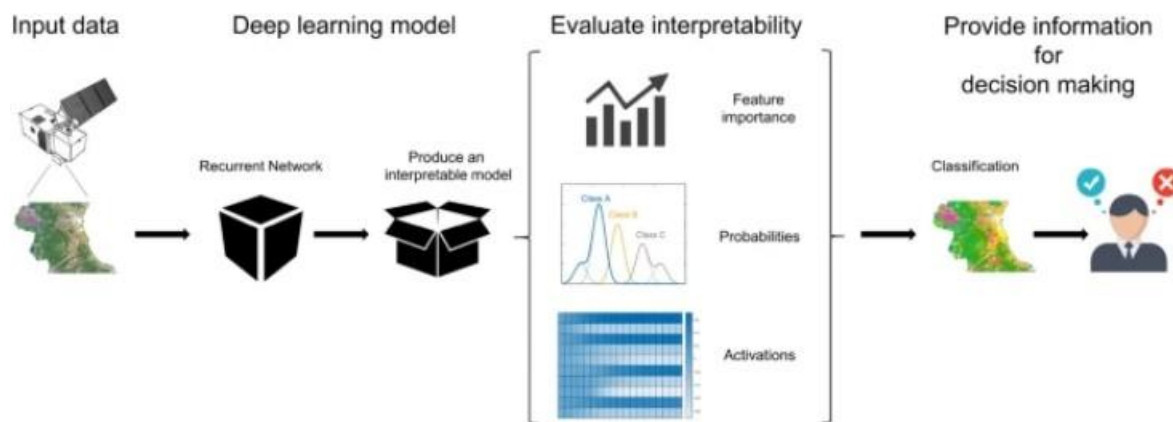


Рисунок 1.3 – Схема запропонованого підходу для поглиблення розуміння рекурентної нейронної мережі для класифікації землекористування на основі даних дистанційного зондування [14]

Таким чином, в рамках міжнародного проекту GEO-GEE українська команда розробила оперативну стійку технологію на основі методів глибокого навчання для класифікації LC/LU, котра в свою чергу базується на супутникових даних, з метою її подальшого використання державою та прийняття рішень. Основна мета полягала у вдосконаленні існуючих методів глибокого навчання для точного картографування LC/LU з використанням супутникових даних з розрізненням 10 м, котрі вже були раніше розроблені та валідовані в межах програми Світового банку.

До того ж було адаптовано та модифіковано методи глибокого навчання для класифікації LC/LU на рівні країни.

Висновки до розділу 1

Проведений аналіз показує, що на даний час розроблено достатньо нових підходів для вдосконалення існуючих глобальних продуктів для більш точного дослідження земельного покриття. Як правило, для цього застосовуються інтелектуальні моделі, зокрема, нейронні мережі. Особливо це стосується локальних досліджень на рівні країни чи навіть менших територіальних одиниць.

Вдосконалено технологію розрахунку індикаторів SDG, яка в поєднанні з моделлю WOFOST дозволяє збільшити просторове розрізнення геопросторових продуктів оцінки індикаторів цілей сталого розвитку.

В описаних вище дослідженнях використовувались хмарні технології та платформи, такі як: Google Earth Engine, Amazon Web Services, Google Cloud Platform, Open Data Cube.

РОЗДІЛ 2 ДЖЕРЕЛА ДАНИХ

Загалом, для класифікації земного покриву використовуються дані, отримані з двох типів джерел: безкоштовні супутникові знімки, отримані в межах програми Copernicus (Sentinel-1 та Sentinel-2), та наземні дані, отримані в результаті наземних досліджень. Може бути використана також будь-яка інша додаткова інформація.

2.1. Супутникові дані

Існує кілька джерел безкоштовних супутникових даних, доступних на регулярній основі, які можна використовувати для картографування LC/LU. Зокрема, українські дослідники здебільшого використовують часові ряди даних Sentinel-1/SAR, отримані з 1 квітня по 15 жовтня за кожний вегетаційний сезон, а також оптичні композити Sentinel-2 [14] за той же період. Цей часовий період вибирається відповідно до календаря посівів для України. Дати, які використовувалися для побудови карти класифікації території України у 2021 році, представлено в Таб. 2.1.

Варто зазначити, що для вирішення задачі класифікації супутник Sentinel-1 надає близько 3 Тб даних для території України, а Sentinel-2 — близько 4 Тб території. У той же час більші країни потребують використання ще більшого об'єму даних для реалізації алгоритму: від 5 до 6 Тб для Sentinel-1 і від 10 до 12 Тб для Sentinel-2 за один вегетаційний рік. Ці величезні обсяги супутникових даних вже зберігаються на Google Earth Engine Platform і Google Cloud Platform. Завдяки цьому

немає необхідності окремого завантаження цих даних, обробки та передачі для подальшого використання.

Таблиця 2.1 – Використані супутникові дані для території України 2021 р.

Satellite data	Dates
Sentinel-2 (composites)	from 01.04.2021 to 31.05.2021;
	from 01.05.2021 to 16.07.2021;
	from 07.07.2021 to 08.08.2021;
	from 08.08.2021 to 02.09.2021;
	from 20.08.2021 to 15.10.2021
Sentinel-1	from 2021-04-01 to 2021-10-15 (with revisit time 6 days for Ukraine)

Охоплення супутниковими даними для України представлено на Рис. 2.1.

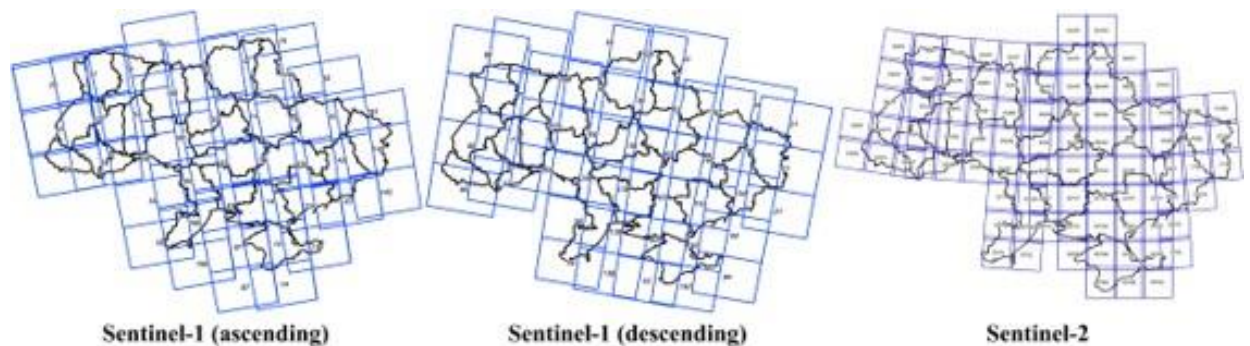


Рисунок 2.1 – Використане для класифікаційного картографування покриття супутниковими даними для України

Інфографіка місії Copernicus Sentinel-1 представлена на Рис. 2.2.

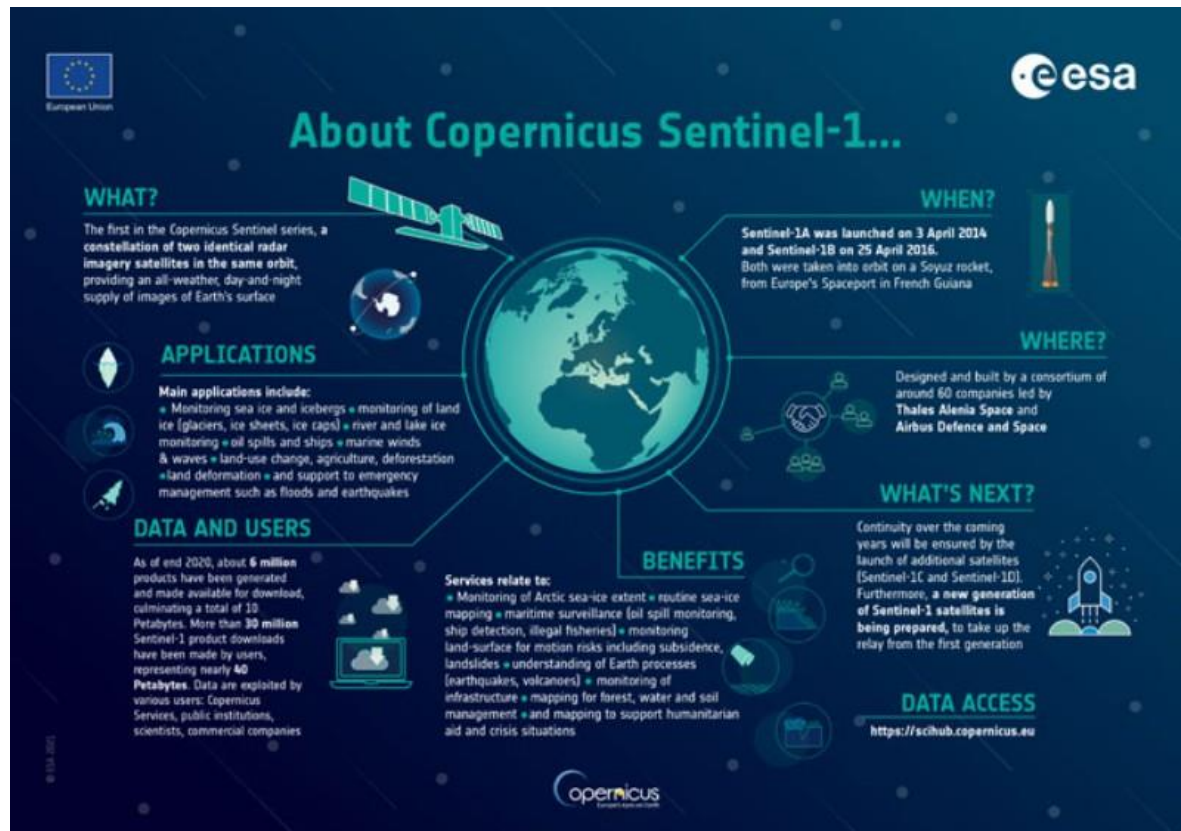


Рисунок 2.2 – Інфографіка місії Copernicus Sentinel-1 [15]

Якщо для дослідження використовується платформа Amazon, то необхідно здійснити такі етапи попередньої обробки супутникових даних SAR:

- 1) читання архіву Sentinel-1 GRD;
- 2) корекція координат на орбіті (app-ply orbit file);
- 3) specl-фільтрація (Lee filter 5 5);
- 4) радіометричне калібрування з перетворенням на значення зворотного коефіцієнта розсіювання сигналу σ_0 (калібрування);
- 5) виконання процедури корекції рельєфу місцевості за даною цифровою моделлю (Shuttle Radar Topography Mission (SRTM) з роздільною здатністю 90 м);

- 6) перетворення даних (LinearToFromdB);
- 7) створення стеку даних (CreateStack);
- 8) збереження смуг зображення (VV, VH);
- 9) злиття смуг VV і VH в межах однієї гранули.

Для оптичних даних використовується атмосферна корекція за допомогою алгоритмів Sen2Cor і маскування хмар та тіней.

2.2 Наземні дані

Збір таких даних проводиться щорічно під час активної стадії вирощування сільськогосподарських культур для отримання достовірної інформації для здійснення якісного навчання нейронної мережі та побудови карти класифікації LC/LU. Планування маршрутів наземної зйомки ґрунтується на основі аналізу супутникових даних часових рядів та офіційних статистичних даних [16]. Маршрут збору даних на місці орієнтований на накопичення максимальної кількості прикладів сільськогосподарських культур, а також на їх максимальну різноманітність. Для цього також були розглянуті дороги, якими можна дістатися до кожного з полів.

Наземні дані отримуються відповідно до рекомендацій JESAM щодо збору даних, наведених у «Керівництві JESAM щодо визначення сільськогосподарських угідь та видів культур та збору наземних даних» [17]. Відповідно до цих рекомендацій було прокладено два маршрути для збору польових даних (на прикладі 2021 року). Один з них - для збору озимих даних (Рис. 2.3, зелена лінія), а інший – для ярих (Рис. 2.3, помаранчева лінія). Вони рівномірно охоплюють

територію України з урахуванням розподілу сільськогосподарських територій та наявності доріг. Маршрут під озимі культури в 2021 році становив близько 1400 км, для літніх — близько 5500 км.

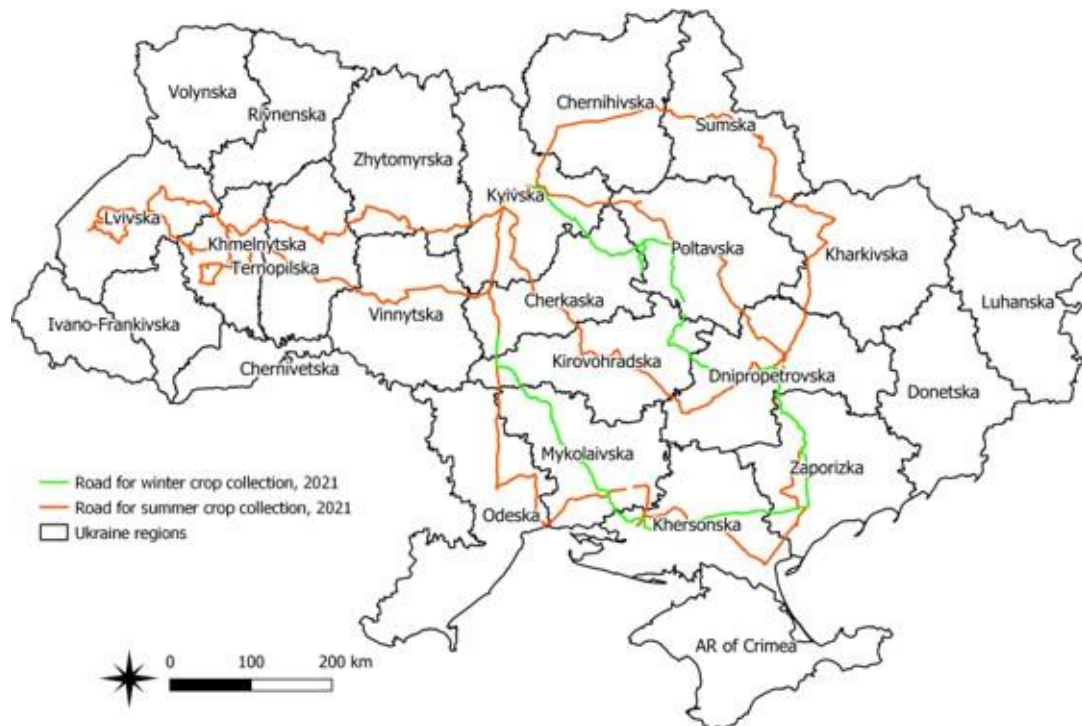


Рисунок 2.3 – Зимові та літні маршрути збору даних на місці у 2021 році

Дані, отримані після етапу попередньої обробки (оцифрування та виявлення аномалій), поділяються на дві незалежні частини для кожної області України та для кожного класу : для навчання нейронної мережі та для тестування отриманої карти класифікації. Детальніше відповідний розподіл за класами та областями можна переглянути в Табл. 2.2.

Таблиця 2.2 – Розподіл даних на місці між тестовими та тренувальними наборами у 2021 році (колір кожної клітинки відповідає класу на карті класифікації)

LC Type (crops)	Train	Test	LC Type	Train	Test
Cereal crops	1858	1855	Artificial	222	223
Rapeseed	241	241	Forest	223	224
Buckwheat	38	38	Grassland	888	866
Maize	915	914	Bare land	135	115
Sugar beet	39	39	Water	351	327
Sunflower	993	992	Wetland	107	114
Soybeans	409	409	Alfalfa	45	34
Other crops	214	205	Gardens	191	176
Peas	63	63	Grape	231	224
Total	4770	4756	Total	2393	2303

Висновки до розділу 2

Для побудови класифікації земного покриття необхідно використовувати супутникові дані та наземні дані. Загалом, ці дані необхідні у доволі значних обсягах, що викликає проблеми з їх обробкою. Цю проблему допомагають вирішити хмарні технології, що дозволяють зберігати дані та користуватись вбудованими функціями для їх аналізу. При використанні AWS супутникові зображення необхідно попередньо обробити за наведеним в пункті 2.1 алгоритмом.

Наземні дані використовуються як для навчання інтелектуальної моделі, так і для оцінки її точності на незалежному наборі даних.

РОЗДІЛ 3 ЗАДАЧА КЛАСИФІКАЦІЇ ТА ОЦІНКА ІНДИКАТОРІВ ЦІЛЕЙ СТАЛОГО РОЗВИТКУ

Передусім хмарні технології полегшують аналіз великих даних і дозволяють краще справлятися із ров'язанням задачі класифікації. У даному розділі розглядається, які саме методи використовуються для класифікації земного покриття на основі раніше отриманих даних та з якими проблемами стикаються дослідники.

Важливо також розглянути методи пошуку оцінки кожного з розглянутих вище індикаторів цілей сталого розвитку і порівняти глобальні продукти з національними розробками, які отримані українськими експертами.

3.1 Класифікація земного покриття

Чимало глобальних продуктів базуються на результатах класифікації земного покриття та інших геопросторових продуктах. Тому дуже важливо отримувати такі продукти класифікації вчасно та з високою точністю на національному рівні. У зв'язку з цим головна ідея всіх українських проектів, які реалізуються на хмарній платформі, – реалізувати кращий підхід до класифікації LC/LU для України. Запропонована методологія заснована на використанні інтелектуальних методів і, зокрема, ансамблю нейронних мереж.

Яскравим прикладом використання такого підходу є стаття [13], в якій було проаналізовано вплив сівоzmіни на біофізичні показники соняшнику на основі

методів регресійного аналізу, а також дослідження [16], де було представлено автоматизований робочий процес класифікації культур з використанням машинного навчання та, власне, його використання на хмарних платформах. Це дозволяє подолати проблеми, пов'язані з великими даними. Діаграма архітектури процесу представлена на Рисунку 3.1.

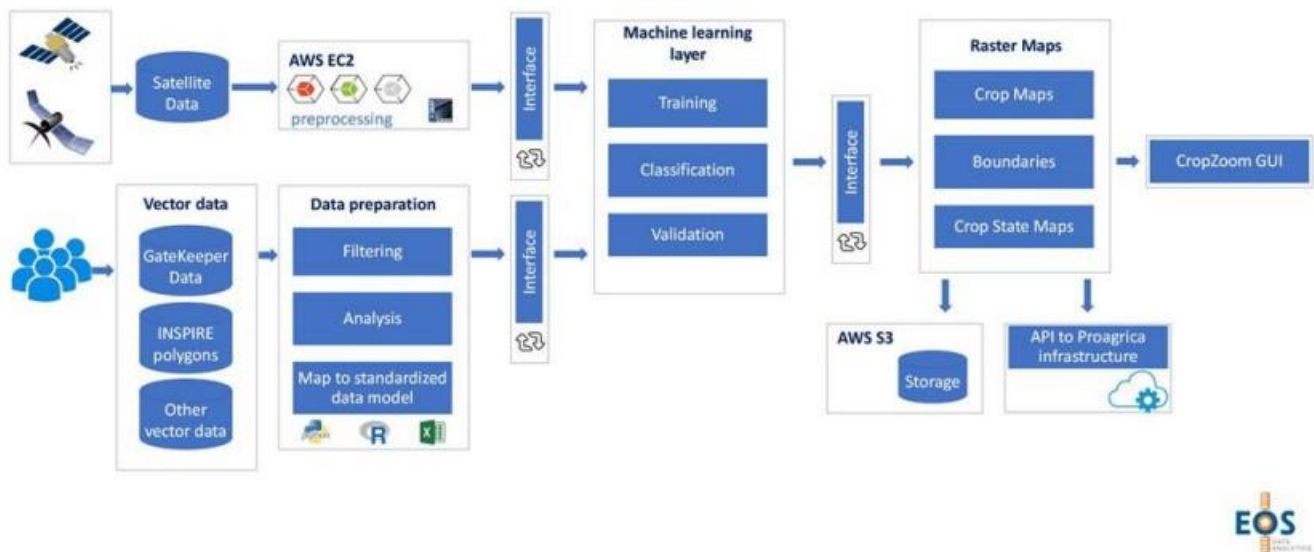


Рисунок 3.1 – Типова діаграма компонентів архітектури для системи CropZoom для класифікації культур [16]

З метою надання результатів класифікації сільськогосподарських культур та карт ґрунтового покриття для конкретної території (починаючи від території України і закінчуючи іншою територією або регіоном) з використанням зображень Sentinel-1 та Sentinel-2 високого просторового розрізнення та наземних даних використано ансамбль комітет нейронних мереж, який за принципом «розділяй та

владарюй» об'єднує вихід декількох нейронних мереж на основі принципу «голосування».

Часові ряди супутникових даних для вегетаційного періоду дозволяють досягти кращої точності класифікації та дозволяють точніше розрізнити сільськогосподарські культури та інші типи ґрунтового покриття. Разом з тим, використання часових рядів супутникових зображень високого просторового розрізнення та їх попередня обробка займає багато часу. Наступна проблема полягає в навчанні моделей Deep Learning на часових рядах супутникових даних.

На даний момент для вирішення наведених вище проблем використання геопросторових даних доступні дві основні потужні хмарні платформи: AWS та Google. Українські експерти мають великий досвід використання потужних хмарних платформ для класифікації LC/LU, таких як Amazon Web Services [20] та безкоштовна інфраструктура Google Earth Engine [21]. Хмарні платформи дозволяють подолати проблеми завантаження та обробки супутникових даних. Платформа Google Earth Engine надає вбудовані функції та доступ до паралельних обчислень. У той же час хмарна платформа Amazon надає можливість використовувати будь-яке програмне забезпечення для обробки зображень, а також бібліотеки з реалізованими сучасними моделями класифікації, зокрема TensorFlow.

Оскільки Amazon не надає готового до використання функціоналу обробки та класифікації супутникових даних, то його використання вимагає значно більших зусиль. Тому був розроблений автоматизований робочий процес для хмарної класифікації культур за допомогою AWS на основі зображень Sentinel-1 та Sentinel-2. Ця методологія була застосована для створення карт класифікації сільськогосподарських культур з просторовим розрізненням 10 м для території України у 2016–2021 рр. Таке вдосконалення дозволило забезпечити набагато кращу точність при оцінці індикаторів цілей сталого розвитку.

Слід підкреслити, що для всіх геопросторових продуктів українські експерти використовували сучасні інформаційні технології з різними видами ресурсів, наприклад, такими як: Elastic Computing 2 (EC2), Simple Storage Service (S3) та інші. Ці інструменти обробки були надані в рамках спільних пілотних проектів комітету GEO та ключових гравців ринку, зокрема компаній Google і Amazon. Для програмування та розробки інтелектуальних моделей реалізовано скрипти Python в середовищі Anaconda та IPython Notebook на основі як стандартних пакетів, так і спеціалізованих, таких як rasterio, gdal, boto3 та інші, що дуже корисні для реалізації моделі, завантаження даних, попередньої обробки, оцінки точності тощо.

3.2 Оцінка індикаторів цілей сталого розвитку

Як випливає з документів Організації Об'єднаних Націй та Паризького саміту (2015 р.), іншим важливим завданням є забезпечення сталого розвитку суспільства. Це можна оцінити на основі використання геопросторової інформації. На основі вже розробленої класифікації LC/LU на рівні країни українська команда запропонувала уніфікований робочий процес (Рис. 3.2) для обчислення індикаторів SDG, а саме індикаторів 11.3.1, 15.1.1, 15.3.1 та 2.4.1 [20, 23]. Запропонований підхід є більш точним, ніж глобальні продукти, та має вище просторове розрішення. Так, у [23] основні вдосконалення полягають у використанні супутникових даних середнього та високого просторового розрізнення та найсучаснішої методології глибокого навчання для класифікації земного покриття та оцінки продуктивності землі.

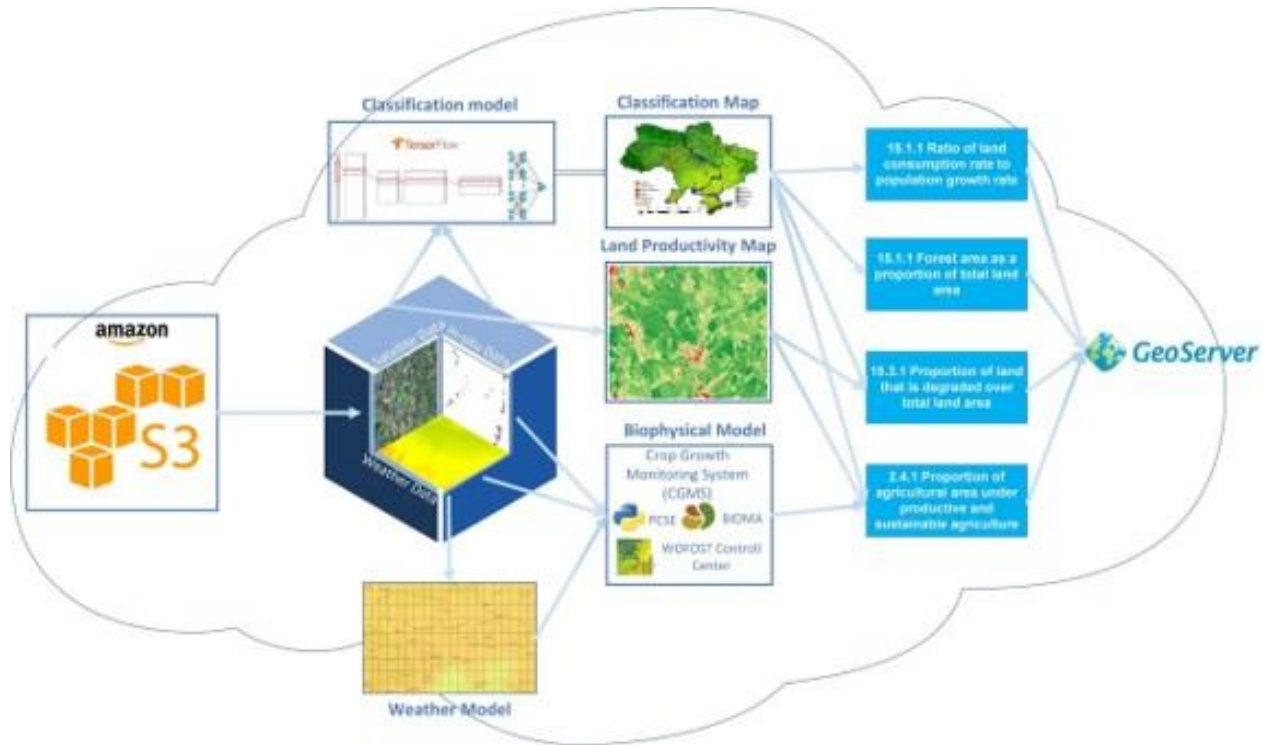


Рисунок 3.2 – Робочий процес для розрахунку показників цілей сталого розвитку 11.3.1, 15.1.1, 15.3.1 та 2.4.1 за допомогою хмарної технології Data Cube

В результаті проведених досліджень було показано, що нові продукти дистанційного зондування з високим просторовим розрізненням можуть суттєво покращити точність оцінювання індикаторів SDG за допомогою спеціальних робочих процесів.

Оцінка всіх індикаторів базується на карті класифікації, допоміжних геопросторових продуктах та реалізованих процедурах інтеграції даних. Деякі з них будуть описані нижче в цьому підрозділі.

3.2.1 Оцінювання індикатора SDG 15.1.1 “Forest area as proportion of total land area”

Індикатори SDG «Площа лісів як частка загальної площі землі» — це частка площ лісів до загальної площі земель. Відповідно до існуючої методології FAO (Food and Agriculture Organization) було зібрано та проаналізовано дані стосовно площі лісів з 1946 року [32]. Це здійснено з інтервалами в 5-10 років у межах програми глобальної оцінки лісових ресурсів (Global Forest Resources Assessment). Продукт FRA 2015 містить інформацію для 234 країн та територій з врахуванням більш ніж 100 особливостей, пов'язаних із площею лісів, їх станом, їх використанням, а також значеннями цих параметрів для п'яти моментів часу: 1990, 2000, 2005, 2010 та 2015 роках. Оцінка площі лісів проводилася через нечасті проміжки часу в багатьох країнах. Варто зазначити, що доступ до зображень дистанційного зондування в останні роки істотно покращився, що дозволяє підвищити якість оцінки індикаторів цілей сталого розвитку.

Іншим доступним глобальним продуктом є Global Forest Change, розроблений Університетом Меріленду з використанням аналізу часових рядів зображень Landsat для характеристики глобальної площі лісу та зміни з 2000 по 2017 рр. з просторовим розрізненням 30 м.

Таким чином, основною необхідною інформацією для оцінки індикаторів є часові ряди карт LC/LU для території, стосовно якої ведеться дослідження. Тому запропоновано використовувати власний національний продукт з 10- метровою класифікацією LC/LU, отриманий за розробленою технологією.

3.2.2 Оцінювання індикатора SDG 15.3.1

Індикатор SDG 15.3.1 дозволяє оцінити частку деградованих земель у загальній площі країни, що базується на аналізі наявних даних або розроблених на національному рівні продуктів, таких як: класифікація земного покриву, продуктивність землі та запаси вуглецю (Табл. 3.1) [24], [25]. Цей індикатор працює за статистичним принципом «One Out, All Out» («один вийшов — усі вийшли») щодо оцінки змін субіндикаторів. Даний принцип означає, що ми маємо три типи змін у субіндикаторах, які зображуються як позитивні чи покращуючі, негативні чи зменшувальні та стійкі чи незмінні.

Таблиця 3.1 – Опис використаних наборів даних

Тип датасету	Джерело	Технічні характеристики	Примітки
Покрив земельних ділянок	ESA CCI land cover dataset	Резолюція 300 м, 1992-2015 рр., 22 класи	22 класи були замінені на 6 основних (ліс, пасовища, посівні угіддя, штучні, водно-болотні угіддя, голі землі). Базовий період 2000-2015 рр. (КБО ООН)
Динаміка продуктивності землі	JRC Productivity Dynamic	роздільна здатність 1 км, 5 класів	На основі часових рядів даних SPOT-Vegetation, зібраних протягом 1999-2013 років доступні 5 якісних класів тенденцій продуктивності землі (зниження продуктивності; ранні ознаки спаду; стабільний, але стресовий; стабільний, не стресовий; підвищення продуктивності)
Запаси органічного вуглецю в ґрунті	ISRIC's SoilGrids 250m	250 метрів	Значення SOC верхнього шару ґрунту (0-30 см)

На жаль, цю методику неможливо використати в повному обсязі, оскільки дуже важко отримати карти запасів вуглецю кожного року на національному рівні [26]. Найновіша карта запасів вуглецю 2016 р. вже є застарілою, а більш сучасні геопросторові продукти на даний момент відсутні.

Якщо один із субіндикаторів має негативні зміни для деякої території, то ця область має негативну продуктивність. Відповідно до існуючої методології, негативними змінами вважаються такі переходи:

- зниження рівня запасів вуглецю протягом певного періоду часу;
- зниження продуктивності землі або негативні зміни ґрунтового покриття (тобто ліси перетворюються в пасовища, ліси перетворюються в угіддя, будь-яка зелена зона стає зоною міського типу).

Власне, підхід, реалізований в Об'єднаному дослідницькому центрі Єврокомісії (Joint Research Centre —JRC), дозволяє врахувати загальну зелену рослинність та є непридатним для оцінки продуктивності посівних угідь, оскільки при цьому найбільш продуктивні регіони часто можуть бути частково вкритими лісами [28]. Крім того, грубе просторове розрізнення цього глобального продукту створює суміші значень пікселів від поєднання різних типів земного покриття в одному пікселі. Тому потрібен власний український продукт земного покриття з високим просторовим розрізненням для більш якісної оцінки даного індикатора, який враховуватиме ці нюанси.

3.2.3 Оцінювання індикатора SDG 11.3.1 “Ratio of land consumption rate to population growth rate”

Індикатор SDG 11.3.1 «Співвідношення темпів споживання землі до темпів зростання населення» [30], [31] можна розрахувати за допомогою детальних карт земного покриття, побудованих на зображеннях середнього та високого просторового розрізнення протягом щонайменше двох років, щоб оцінити співвідношення коефіцієнта споживання землі. При цьому необхідно також мати відкритий доступ до статистичних даних про населення міста. Для оцінки темпів зростання населення використовується статистика населення міста, надана органами статистики. У випадку, якщо національні дані недоступні, можна використовувати відкриті дані, такі як глобальний продукт JRC (JRC Global Human Settlement) [32].

Цей індикатор можна використовувати в кожній країні світу, якщо користуватись глобальною картою земного покриття низького просторового розрізнення. Інформацію можна отримати на двох рівнях:

- 1) Перший рівень – індикатор на рівні міста, який вимірює стійкість із точки зору міста та зростання населення.
- 2) Другий рівень — індикатор на рівні країни, який агрегує показники за містами в межах країни та вказує на територію міст і темпи зростання населення для цієї країни.

3.2.4 Оцінювання індикатора SDG 2.4.1 “Proportion of agricultural area under productive and sustainable agriculture”

Робочий процес розрахунку індикатора 2.4.1 SDG «Пропорція сільськогосподарської площі під продуктивним та сталим сільським господарством» [26] включає визначення продуктивності землі, зокрема над посівними угіддями. Методологія, запроваджена JRC, враховує загальну зелену рослинність і непридатна для оцінки продуктивності сільськогосподарських угідь, оскільки найбільш продуктивні регіони часто покриті лісами [28]. Крім того, низьке просторове розрізнення цього глобального продукту зумовлює наявність сумішей значень пікселів від поєднання різних типів земного покриття в одному пікселі.

Індикатор SDG 2.4.1 може бути розрахований з використанням тієї ж методології, що і для розрахунку індикатора 15.3.1. Цей показник являє собою частку сільськогосподарської площі, яка має позитивне значення тенденції продуктивності до загальної сільськогосподарської площі за таким же правилом «Один вийшов — усі вийшли». Для цього індикатора використовуються ті самі субіндикатори, що й для індикатора 15.3.1, але територією інтересу є не вся площа країни, а сільськогосподарські угіддя (посівні угіддя), згруповані за допомогою карти земного покриття. Також у цьому випадку використання супутникових зображень високого просторового розрізнення є особливо важливим, оскільки змішані пікселі суттєво впливають на величину змін субіндикатора.

Висновки до розділу 3

З метою пошуку кращого підходу до побудови LC/LU карт та оцінки індикаторів SGD для України було запропоновано кілька нових підходів, таких як: автоматизація робочого процесу класифікації культур з використанням хмарних технологій та AWS зокрема; використання ансамблю нейронних мереж; використання часових рядів спутникових даних; використання субіндикаторів для оцінки індикатора.

Тим не менше, потрібне вдосконалення або пошук нових підходів для оцінки індикатора 15.3.1 через проблеми з доступом або наявністю актуальних вуглецевих карт.

РОЗДІЛ 4 РЕЗУЛЬТАТИ РОЗВ’ЯЗАНИХ ПРИКЛАДНИХ ЗАДАЧ

Для того, щоб чітко розуміти якість наявних методів та підходів і використовувати їх в подальшому для аналізу земельного покриву та оцінки індикаторів SDG, необхідно мати уявлення про точність тих чи інших методів та значень індикаторів.

4.1 Результати побудови карт класифікації

Як вже зазначалося, основною передумовою оцінки майже всіх індикаторів SDG є карта класифікації. Тому дуже важливо оцінити точність класифікації. З використанням розробленого підходу для основних сільськогосподарських культур України у 2021 році (Рис. 4.1) загальна точність класифікації земельного покриву була вище 94%.

Матриці похибок для класів було побудовано на основі незалежних тестових даних для карти, що була отримана за допомогою методу Random Forest в Google Earth Engine та з використанням нейронних мереж у середовищі Amazon Web Services.

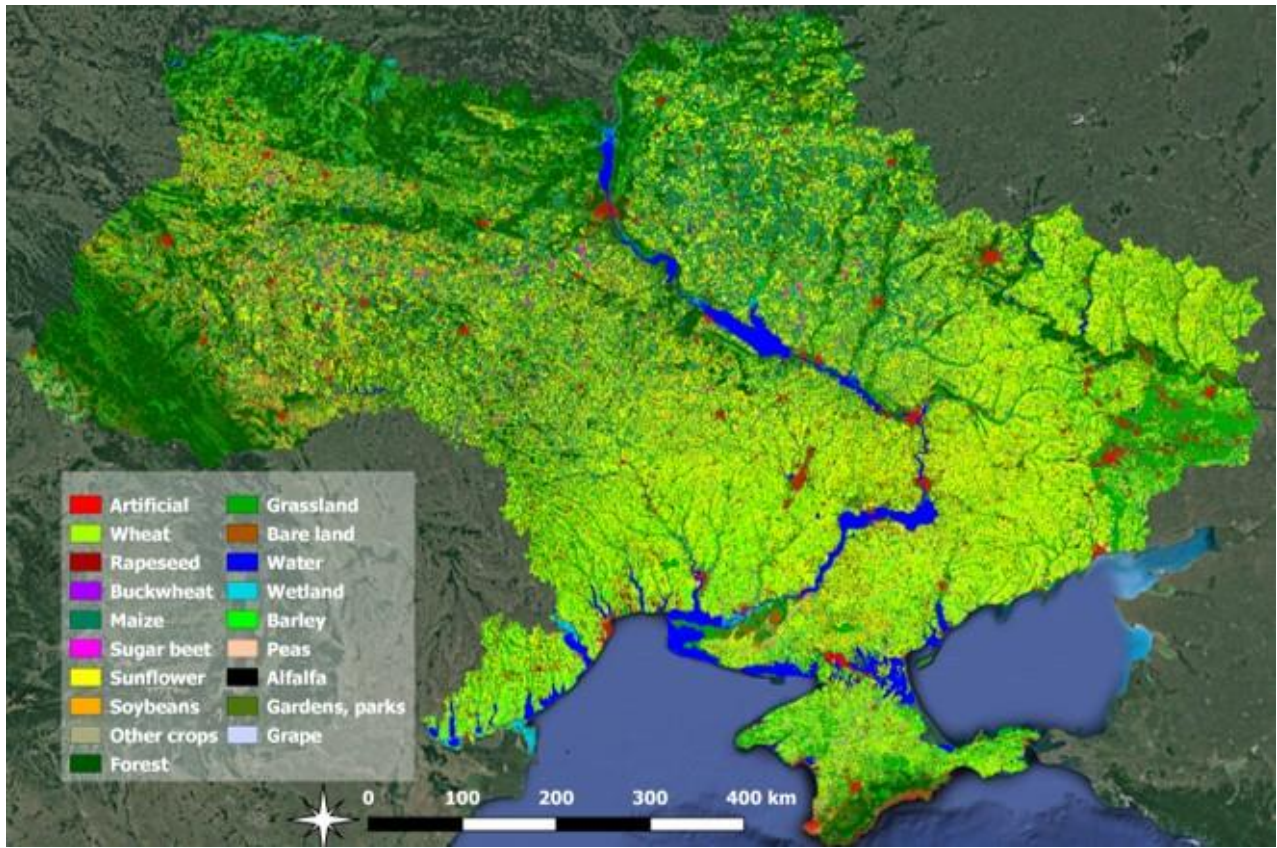


Рисунок 4.1 – Класифікаційна карта для України, 2021 рік

Вищезгадані матриці похибок представлено на Рис. 4.2 та Рис. 4.3. За допомогою методу Random Forest загальна точність моделі склала 94,8%, а загальна точність моделі, побудованої шляхом використання нейромережевої моделі, досягла 96,7% [24], що є дуже хорошим результатом.

Kappa = 0.94	Artificial	Cereal crops	Rapeseed	Buckwheat	Maize	Sugar beet	Sunflower	Soy	Other crops	Forest	Grassland	Bare land	Water	Wetland	Peas	Alfalfa	Gardens, parks	Grape	UA	F1
Artificial	22567	0	0	0	3	0	0	0	16	0	42	12561	4053	39	0	0	26	220	57,1	70,6
Cereal crops	294	1865708	1355	3795	3548	0	1614	98	5182	0	10318	219	263	36	26334	5729	7926	4422	96,3	97,8
Rapeseed	0	0	464897	5330	0	0	0	0	114	0	0	0	0	0	2734	0	0	630	98,1	98,9
Buckwheat	0	0	0	1823	0	0	0	0	0	0	0	0	0	0	0	0	0	0	100	24
Maize	10	175	0	163	562293	1546	2469	14863	1008	0	647	9	17	516	173	20	2641	1259	95,7	96,9
Sugar beet	0	1	0	0	136	32510	270	1326	645	0	0	0	0	9	3	0	1	0	93,1	86,6
Sunflower	18	412	0	762	712	4301	792949	10824	2246	0	2	76	0	691	13	340	322	97	97,4	98,2
Soy	0	2662	0	741	4726	1732	2947	143695	1441	0	158	4	0	175	138	1	989	5	90,1	86,9
Other crops	170	397	88	3	82	78	310	14	4796	0	80	374	3550	18	56	0	676	390	43,3	34,4
Forest	439	0	0	0	4	0	0	1	9	87527	43	0	13	427	0	0	16261	0	83,6	91
Grassland	234	8558	45	536	1233	0	450	167	580	23	163312	130	312	6380	984	11268	29861	4756	71,4	80,5
Bare land	404	52	0	0	36	0	0	7	1	0	56	16558	4359	4	12	0	268	3	76,1	64
Water	14	0	0	0	0	0	0	0	0	1	4	773905	114	0	0	0	0	100	99,2	
Wetland	50	4	0	0	135	0	0	23	12	0	411	3	41	8682	0	0	39	4	92,3	63,5
Peas	0	1	31	0	0	0	0	0	472	0	0	0	0	0	33884	0	0	0	98,5	68,6
Alfalfa	0	531	0	0	0	0	0	0	0	0	15	0	0	9	12	19662	0	0	97,2	68,5
Gardens, parks	166	54	0	197	94	0	184	104	160	192	1746	34	14	820	0	132	54795	13261	76,2	59
Grape	32	1	0	4	0	0	10	0	123	0	3	15	5	0	0	0	144	11059	97	46,6
PA	92,5	99,3	99,7	13,7	98,1	80,9	99	84	28,5	99,8	92,4	55,2	98,4	48,4	52,7	52,9	48,1	30,6	84,8	

Рисунок 4.2 – Матриця похибок для основних класів, отриманих на тестовому наборі даних на місці для методу Random Forest

Kappa = 0.96	Artificial	Cereal crops	Rapeseed	Buckwheat	Maize	Sugar beet	Sunflower	Soy	Other crops	Forest	Grassland	Bare land	Water	Wetland	Peas	Alfalfa	Gardens, parks	Grape	UA	F1
Artificial	21763	139	4	0	0	0	28	5	55	187	653	3860	2183	54	0	5	963	46	72,7	80,1
Cereal crops	5	1854789	1063	12	1096	0	5	26	228	0	2744	3	21	331	1818	920	659	4	99,5	99,1
Rapeseed	1	1139	464795	0	15	0	0	0	491	0	474	0	16	0	226	6	2	0	99,5	99,6
Buckwheat	2	167	22	13084	702	20	617	3451	127	0	192	0	0	39	604	0	0	0	68,8	80,8
Maize	2	2	0	0	558356	92	2646	2683	290	7	100	1	331	796	0	144	38	0	98,7	98,1
Sugar beet	8	0	0	124	1109	39840	3848	311	68	0	610	0	9	238	0	0	54	71	86,1	92,2
Sunflower	10	728	0	4	3764	62	789533	2145	630	0	334	169	218	154	74	0	79	7	99	98,8
Soy	0	227	0	0	6636	134	673	160965	171	0	365	0	1	122	0	13	163	0	95	94,5
Other crops	2418	8118	372	51	370	19	1707	708	12118	0	1030	6610	1491	493	93	1807	377	166	31,9	44,3
Forest	0	0	0	0	1	0	0	0	213	86488	65	0	3	355	0	0	2718	0	96,3	97,4
Grassland	71	3588	99	9	83	0	459	24	717	0	128295	9	191	3845	0	558	8409	169	87,6	79,4
Bare land	98	0	0	0	0	0	0	0	14	0	80	19269	10894	0	0	0	0	0	63,5	63,9
Water	15	0	0	0	0	0	0	0	0	0	0	62	771007	17	0	0	0	0	100	99
Wetland	0	0	0	0	111	0	6	1	0	0	1152	0	106	10546	0	0	0	0	88,5	70,8
Peas	0	1802	4	41	56	0	100	354	891	0	32	0	0	0	61510	1094	5	0	93,4	94,5
Alfalfa	0	1610	1	0	77	0	0	24	56	0	1451	0	11	1	0	32135	136	0	90,5	88,4
Gardens, parks	10	6162	52	25	695	0	1035	346	639	1071	38255	0	35	891	1	394	98746	1666	65,8	74,8
Grape	0	1	0	0	4	7	454	7	88	0	952	2	3	2	0	87	1537	33974	91,5	92,8
PA	89,2	98,7	99,7	98	97,4	99,2	98,6	94,1	72,1	98,6	72,6	64,3	98	59	95,6	86,5	86,7	94,1	86,7	

Рисунок 4.3 – Матриця похибок для основних класів, побудованих на тестовому наборі даних для методу на основі нейронних мереж.

Обидві карти класифікації порівнювалися з офіційною статистикою, отриманою від Державної служби статистики України для різних культур. Наприклад, для Київської області України загальна точність карти класифікації покращена більш ніж на 10% порівняно з набором даних ESA Climate Change Initiative Land Cover.

Розглянемо більш детально коефіцієнт каппа, який показує, наскільки дані, зібрані під час наземних досліджень, є правильними представленнями вимірних змінних. Для набору даних ESA Climate Change Initiative Land Cover цей коефіцієнт становить 0,75, тоді як коефіцієнт каппа для карти, побудованої українською групою експертів, становить 0,9.

Порівняння ґрунтового покриття ESA CCI із ґрунтовим покритвом Київської області на основі підходу з використанням нейронних мереж показано на Рис. 4.4. Як можна побачити, наведені в цілому відрізняються лише інтенсивністю параметра «Grassland», що відповідає пасовищам. Тобто, у дослідженні, проведеному ESA CCI, було виявлено значне завищення площ сільськогосподарських угідь.

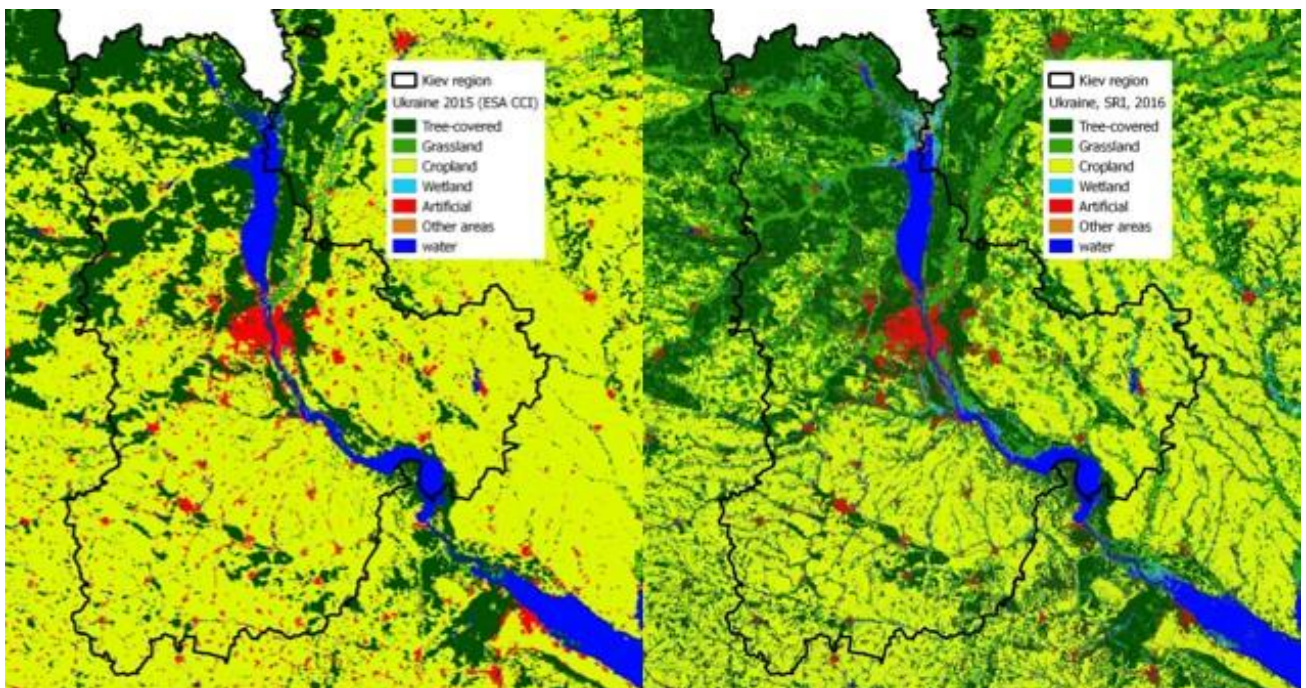


Рисунок 4.4 – Порівняння ґрунтового покриття ESA CCI (ліворуч) із земельним покритвом (праворуч) над Київською областю

Результати класифікації для Київської області, що були отримані в Google Cloud Platform, представлені на Рис. 4.5.

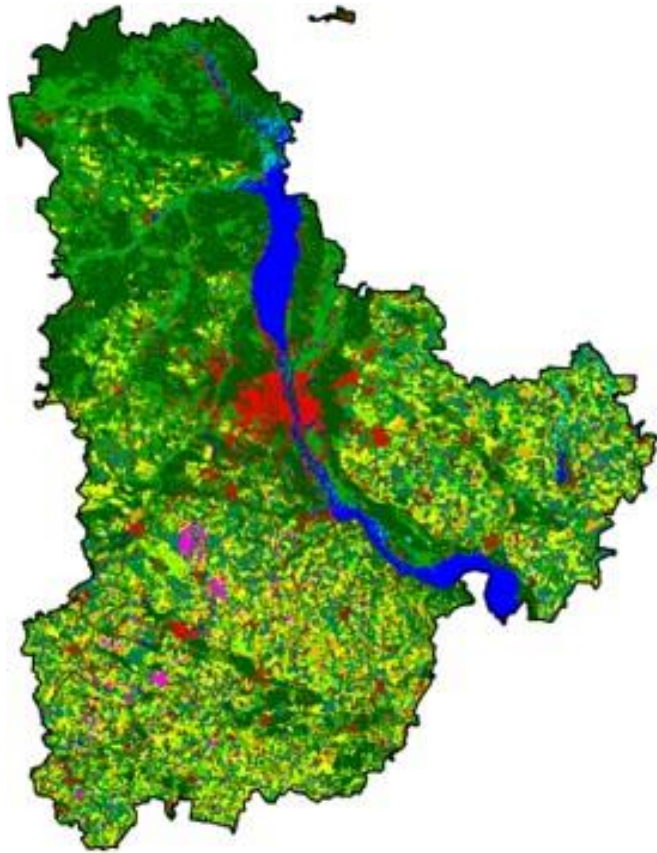


Рисунок 4.5 – Результати класифікації для Київської області в Google Cloud Platform

Варто також відзначити, яка саме відмінність у точності спостерігалась в залежності від використовуваних методів та платформ для розрахунків та побудови карт класифікації. Загальна точність такої карти, отриманої шляхом використання нейромережевої моделі та Google Cloud Platform, становить 95%, що на 2% вище, ніж точність карти класифікації, яка побудована з використанням Random Forest та GEE. Детальніше для кожного класу культур ці значення можна переглянути в

Табл. 4.1. Відмінності між такими картами особливо помітні на міноритарних (малочисельних) посівах, для яких використання Random Forest дає значно гірші результати.

Таблиця 4.1 – Порівняння точності алгоритмів NN і RF в Google Cloud Platform і Google Earth Engine

	Google Cloud Platform (NN)			GEE (RF)		
	UA	PA	F1	UA	PA	F1
Artificial	92,6	85,5	88,9	82,3	81,1	81,7
Wheat	90,1	96,9	93,4	84,4	98,3	90,8
Barley	86,9	49,3	62,9	47,6	7,4	12,7
Rape	100	100	100	99,1	99,8	99,4
Maize	97,5	97,4	97,4	97,1	95,3	96,2
Sugar beet	99,6	99,9	99,8	99,5	98,7	99,1
Sunflower	96,9	98,7	97,8	96,3	95,4	95,9
Soy	94,2	89,6	91,9	87,6	90,2	88,9
Other crops	55,8	64,4	59,8	10,6	8,5	9,5
Forest	99,8	99,7	99,8	98,8	99,8	99,3
Grassland	86	97,7	91,5	78,1	94,7	85,6
Bare land	71,6	70,9	71,2	74,5	72	73,3
Water	99,9	100	100	100	100	100
Wetland	76,5	89,9	82,7	95,7	30,6	46,3
Peas	100	89	94,2	100	31,6	48
Overall accuracy	95,8			93,1		

У майбутньому доцільно використовувати Google Cloud buckets із супутниковими даними та створювати карти LC/LU для всієї території України методами глибокого навчання. Таким чином, згодом карта ґрунтового покриття зможе вирішити проблеми прозорого землекористування, реалізації аграрної політики та сталого розвитку суспільства в Україні.

4.2 Результати оцінки індикаторів SDG

Використовуючи карти земного покриття за 2016 та 2019 роки з просторовим розрізненням 10 м, було обчислено зміну забудови для міст України та з допомогою продукту NASA Gridded Population of the World отримано індикатор SDG 11.3.1 у масштабі міста та країни (Рис. 4.6). Щоб провести розрахунок цього індикатора для України, було створено Functional Urban Area layer, сумісний із продуктами Copernicus Urban Atlas [35].

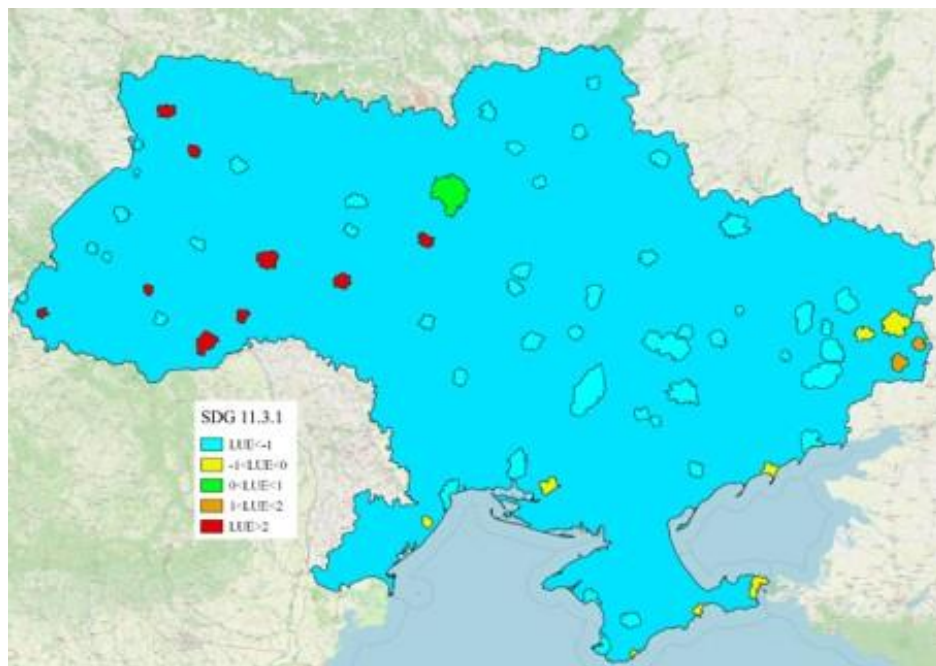


Рисунок 4.6 – Індикатор SDG 11.3.1 «Ratio of land consumption rate to population growth rate» для території України

Індикатор SDG 15.1.1 за 2019–2021 рр. для окремих регіонів та України в цілому оцінено за допомогою карти земного покриття (Рис. 4.7). Отримані

результати підтверджуються наявною інформацією за даними Державної служби статистики України. Середнє відхилення розрахункових значень від статистичних показників склало близько 5%.

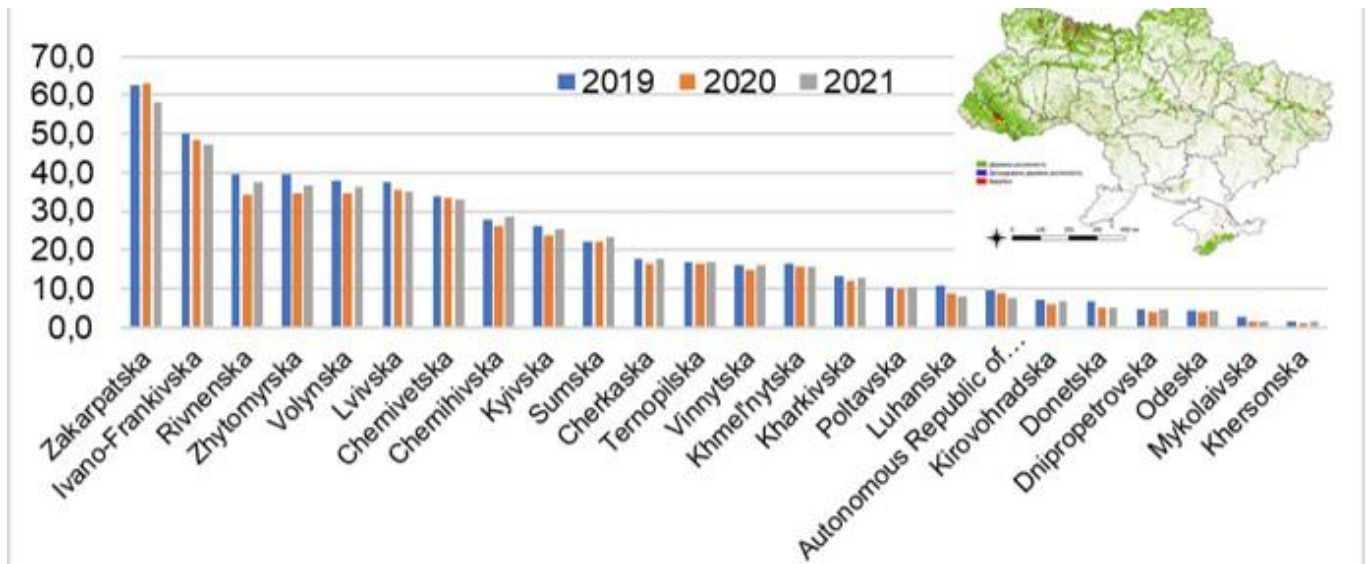


Рисунок 4.7 – Індикатор SDG 15.1.1 «Forest area as proportion of total land area» для території України

І, нарешті, як іще один цінний продукт, який базується на класифікаційних картах, створено попередню карту деградації земель для всієї України з просторовим розрізненням 30 м. Для цього використано аналіз трендів індексів рослинності супутникових даних Landsat-8 та Sentinel-2. Цей продукт є основою для розрахунку індикаторів SDG 2.4.1 та 15.3.1, що в процесі його створення є найбільш трудомістким завданням (Рис. 4.8).

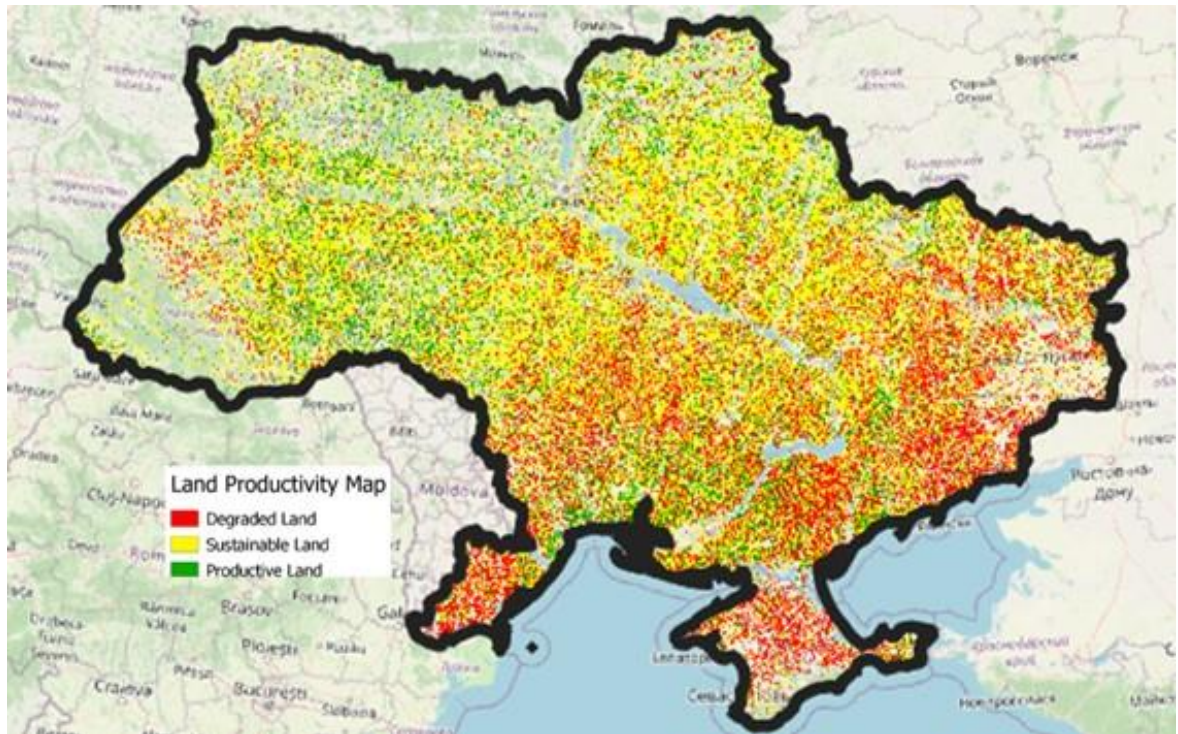


Рисунок 4.8 – Карта продуктивності землі для території України

Висновки до розділу 4

Таким чином, у питанні створення LC/LU карт глобальні продукти програють у точності національним. При цьому кращі результати було отримано на платформі Google Cloud з використанням нейромережевих моделей. Також було отримано оцінку індикаторів SDG 11.3.1 (використовуючи Functional Urban Area layer), 15.1.1 (на основі карти класифікації земного покриття), 15.3.1 та 2.4.1 (базуються на картах деградації землі).

ВИСНОВКИ

При виконанні досліджень було проаналізовано інтелектуальні моделі, які на даний момент використовуються для розв'язання прикладних задач класифікації земного покриття та оцінки індикаторів цілей сталого розвитку, а також хмарні технології, що використовуються для класифікації земельного покриття та оцінювання індикаторів сталого розвитку SDG в Україні. Здійснено аналіз існуючих дослідницьких праць на дану тематику. Враховуючи, що для створення LC/LU карт здебільшого необхідні супутникові дані та наземні дані за вегетаційний період, хмарні технології є потужним засобом, що значно полегшує роботу з подібними обсягами інформації. Українські експерти мають доступ до ряду інструментів, призначених для обробки великих даних, що були надані в межах спільних пілотних проєктів комітету GEO та ключових гравців ринку, зокрема компаній Google і Amazon, а саме Amazon Web Services та Google Earth Engine.

Усі геопросторові продукти для оцінки індикаторів SDG базуються на ряді основних геопросторових продуктів, а саме на картах класифікації за різні роки [35], DEM [36], векторних даних з OpenStreetData. Окрім супутникових даних, які доступні в хмарних інфраструктурах і готові до використання, сформовано також і наземні дані, які використовувались для формування навчальних та тестових вибірок роботи з інтелектуальними моделями [37]. Для карти класифікації загальна точність становила близько 95%, тоді як для всіх класів точність становила від 75% до 98%. На основі отриманих результатів в подальшому можна генерувати геопросторові продукти на рівні країни, які є більш точними, ніж на основі глобальних або регіональних даних із низьким просторовим розрізненням. Зокрема, на основі національних карт можна створювати різні геопросторові

продукти для оцінки індикаторів SDG, карти деградації [38], карти продуктивності [39], продукти оцінки якості повітря [40] та вирішувати інші важливі прикладні завдання. Також, використовуючи розроблені карти [41], можна побудувати додаткові продукти, наприклад, для оцінки рівнів деградації земель із просторовим розрізненням 10 м [42].

Враховуючи, що для отримання геопросторових продуктів пропонується використовувати дані з просторовим розрізненням 10 м, питання оптимізації часових рядів та підвищення ефективності їх обробки варто також досліджувати в майбутньому.

ПЕРЕЛІК ВИКОРИСТАНИХ ДЖЕРЕЛ

1. The World Bank: Projects [Електронний ресурс]. – Режим доступу: <https://projects.worldbank.org/en/projects-operations/projects-home>. — Назва з екрану. — Дата звернення: 26.05.2022.

2. Закон України "Про внесення змін до деяких законодавчих актів України щодо умов обігу земель сільськогосподарського призначення"[Електронний ресурс]. – Режим доступу: <https://zakon.rada.gov.ua/laws/show/552-20#Text>. – Назва з екрану. — Дата звернення: 26.05.2022.

3. Національна доповідь. Цілі Сталого Розвитку: Україна [Електронний ресурс]. – Режим доступу: https://mepr.gov.ua/files/docs/%D0%9D%D0%B0%D1%86%D1%96%D0%BE%D0%BD%D0%B0%D0%BB%D1%8C%D0%BD%D0%B0%20%D0%B4%D0%BE%D0%BF%D0%BE%D0%B2%D1%96%D0%B4%D1%8C%20%D0%A6%D0%A1%D0%A0%20%D0%A3%D0%BA%D1%80%D0%B0%D1%97%D0%BD%D0%B8_%D0%BB%D0%B8%D0%BF%D0%B5%D0%BD%D1%8C%202017%20ukr.pdf Назва з екрану. — Дата звернення: 26.05.2022.

4. Belgiu, M., and Ovidiu C.: Sentinel-2 cropland mapping using pixel- based and ob-ject-based time-weighted dynamic time warping analysis. Remote sensing of envi-ronment 204, pp. 509–523, (2018).

5. Menno Jan, K., Ricker B., Engelhardt Y.: Challenges of mapping sustainable de-velopment goals indicators data. ISPRS international journal of geo-information 7.12: 482, (2018).

6. Compendium of Earth Observation contributions to the SDG Targets and Indicators [Electronic Resource]. – Mode of access: https://eo4society.esa.int/wp-content/uploads/2021/01/EO_Compendium-for-SDGs.pdf, last accessed 26.05.2022

7. Open Data Cube (ODC) [Electronic Resource]. – Mode of access: https://eo4society.esa.int/wp-content/uploads/2021/01/EO_Compendium-for-SDGs.pdf – Title from the screen. – Last accessed 26.05.2022

8. Diepen, C.A., Rappoldt, C., Wolf, J., Keulen, H.: Crop growth simulation model WOFOST. Documentation version 4.1. for World Food Studies. Wageningen, The Netherlands. (1988).

9. Food and agriculture organization of the United Nations. Indicator 2.4.1 - Proportion of agricultural area under productive and sustainable agriculture [Electronic Resource]. – Mode of access: <https://www.fao.org/sustainable-development-goals/indicators/241/en/> – Title from the screen. – Last accessed 20.05.2022

10. Indicator 11.3.1 [Electronic Resource]. – Mode of access: <https://unstats.un.org/sdgs/metadata/files/Metadata-11-03-01.pdf> – Title from the screen. – Last accessed 26.05.2022

11. Indicator 15.1.1 [Electronic Resource]. – Mode of access: <http://www.fao.org/sustainable-development-goals/indicators/1511/en/> – Title from the screen. – Last accessed 26.05.2022

12. Indicator 15.3.1 [Electronic Resource]. – Mode of access: <https://knowledge.unccd.int/topics/sustainable-development-goals-sdgs/sdg-indicator-1531> – Title from the screen. – Last accessed 26.05.2022

13. Kussul, N., Lavreniuk, M., Shumilo L.: Deep recurrent neural network for crop classification task based on Sentinel-1 and Sentinel-2 imagery. In IGARSS 2020 IEEE

International Geoscience and Remote Sensing Symposium. Waikoloa, HI, USA. P. 6914–6917. DOI: 10.1109/IGARSS39084.2020.9324699 (2020).

14. Copernicus Sentinel mission [Electronic Resource]. – Mode of access: <http://copernicus.eu/main/sentinels> – Title from the screen. – Last accessed 26.05.2022

15. Sentinel-1 [Electronic Resource]. – Mode of access: <https://sentinel.esa.int/web/sentinel/missions/sentinel-1> – Title from the screen. – Last accessed 26.05.2022

16. State Statistics Service of Ukraine [Electronic Resource]. – Mode of access: <http://www.ukrstat.gov.ua/> – Title from the screen. – Last accessed 26.05.2022

17. Understanding deep learning in land use classification based on Sentinel-2 time series [Electronic Resource]. – Mode of access: <https://www.nature.com/articles/s41598-020-74215-5> – Title from the screen. – Last accessed 26.05.2022

18. JECAM Guidelines for cropland and crop type definition and field data collection [Electronic Resource]. – Mode of access: http://jecam.org/wp-content/uploads/2018/10/JECAM_Guidelines_for_Field_Data_Collection_v1_0.pdf – Title from the screen. – Last accessed 26.05.2022

19. Shelestov, A., Lavreniuk, M., Vasiliev, V., Shumilo, L., Kolotii, A., Yailymov, B., Kussul, N., Yailymova, H.: Cloud Approach to Automated Crop Classification Using Sentinel-1 Imagery IEEE Transactions on Big Data. Vol. 6, No. 3. pp. 572–582. DOI: 10.1109/TBDATA.2019.2940237 (2020).

20. Shelestov, A., Lavreniuk, M., Kussul, N., Novikov, A., Skakun, S.: Exploring Google earth engine platform for Big Data Processing: Classification of multi-temporal satellite imagery for crop mapping. Frontiers in Earth Science, 5, 17 (2017).

21. Shelestov, A., Lavreniuk, M., Kolotii, A., Vasiliev, V., Shumilo, L., Kussul, N.: Cloud approach to automated crop classification using Sentinel-1 imagery. Proc. of the 2017 conference on Big Data from Space (BiDS'17). pp. 122–125 (2017).
22. Sims, Neil C., et al.: A land degradation interpretation matrix for reporting on UN SDG indicator 15.3.1 and land degradation neutrality. *Environmental Science and Policy* 114: pp. 1–6, (2020).
23. Kussul, N., Lavreniuk, M., Kolotii, A., Skakun, S., Rakoid, O., Shumilo, L.: A workflow for Sustainable Development Goals indicators assessment based on high-resolution satellite data. *International Journal of Digital Earth*. Vol. 2, No. 13. pp. 309–321. DOI: 10.1080/17538947.2019.1610807 (2020).
24. Kaifeng, P., et al.: Evaluating the potential impacts of land use changes on ecosystem service value under multiple scenarios in support of SDG reporting: A case study of the Wuhan urban agglomeration. *Journal of Cleaner Production* 307: 127321 (2021).
25. Rizvi, Syed R., et al.: Data Cube Application Algorithms for the United Nations Sustainable Development Goals (UN-SDGs). In *IGARSS 2020 IEEE International Geoscience and Remote Sensing Symposium*. IEEE, 2020.
26. Hamilton, Stuart E., and Friess D.: Global carbon stocks and potential emissions due to mangrove deforestation from 2000 to 2012. *Nature Climate Change* 8.3: 240–244 (2018).
27. Kussul, N., et al.: Nexus Approach for Calculating SDG Indicator 2.4.1 Using Remote Sensing and Biophysical Modeling. In *IGARSS 2019 IEEE International Geoscience and Remote Sensing Symposium*. IEEE, 2019.
28. Barbero Vignola, G., S. Acs, and S. Borchardt: Modelling for Sustainable Development Goals (SDGs): Overview of JRC models. (2020).

29. Kussul, N., Shelestov, A., Yailymova, H., Yailymov, B., Lavreniuk, M., Ilyashenko, M.: Satellite agricultural monitoring in Ukraine at country level: world bank project. In IGARSS 2020-2020 IEEE International Geoscience and Remote Sensing Symposium, pp. 1050–1053 (2020).

30. Kussul, N., Shelestov, A., Lavreniuk, M., Yailymov B., Kolotii A., Yailymova H., Skakun S., Shumilo L., Bilokonska Y.: SDG indicator 11.3.1 within HORIZON-2020 SMURBS, Space research in Ukraine. 2018–2020. pp. 91–95 (2020).

31. Shelestov, A., et al. Assessment of land consumption for sdg indicator 11.3.1 using global and local built-up area maps. In IGARSS 2020-2020 IEEE International Geoscience and Remote Sensing Symposium (2020).

32. Global Human Settlement layer [Electronic Resource]. – Mode of access: <https://ghsl.jrc.ec.europa.eu/datasets.php> – Title from the screen. – Last accessed 25.05.2022

33. Forest resources assessment [Electronic Resource]. – Mode of access: <http://www.fao.org/forest-resources-assessment/en/> – Title from the screen. – Last accessed 25.05.2022

34. World Soil Information [Electronic Resource]. – Mode of access: <https://www.isric.org/> – Title from the screen. – Last accessed 25.05.2022

35. Shumilo, L., et al.: Is Soil Bonitet an Adequate Indicator for Agricultural Land Appraisal in Ukraine. Sustainability 13.21: 12096 (2021).

36. Liu, Xiaoye. Airborne LiDAR for DEM generation: some critical issues. Progress in physical Geography 32.1: 31–49 (2008).

37. Kussul, N., et al. Crop state and area estimation in Ukraine based on remote and in-situ observations. Int. J. on Information Models and Analyses 1.3: 251–259 (2012).

38. Kussul, N., et al. Land degradation estimation from global and national satellite based datasets within UN program. In 2017 9th IEEE International Conference on Intelligent Data Acquisition and Advanced Computing Systems: Technology and Applications (IDAACS). Vol. 1 (2017).

39. Kussul, N., et al. Cropland productivity assessment for Ukraine based on time series of optical satellite images. In 2017 IEEE International Geoscience and Remote Sensing Symposium (IGARSS) (2017).

40. Shelestov, A., et al. Air Quality Estimation in Ukraine Using SDG 11.6. 2 Indicator Assessment. Remote Sensing 13.23: 4769 (2021).

41. N. Kussul, A. Shelestov, M. Lavreniuk, I. Butko and S. Skakun, "Deep learning approach for large scale land cover mapping based on remote sensing data fusion," 2016 IEEE International Geoscience and Remote Sensing Symposium (IGARSS), 2016, pp. 198–201.

42. N. Kussul, N. Lavreniuk, A. Shelestov, B. Yailymov, and I. Butko Land Cover Changes Analysis Based on Deep Machine Learning Technique.- Journal of Automation and Information Sciences. — 2016. — PP. 42–54.

43. V.N Azarskov, L.N. Blohin, L.S. Zhitetskii, and N.N. Kussul (2004). Robust methods for estimation, identification and adaptive control. — K.: NAU. — 500 p.