

Національна академія наук України
Державне космічне агентство України
Інститут космічних досліджень

На правах рукопису

Яйлимов Богдан Ялкапович

УДК 528.8.04:004.021

**АВТОМАТИЗОВАНА ІНФОРМАЦІЙНА ТЕХНОЛОГІЯ КАРТОГРАФУВАННЯ ЗЕМНОГО
ПОКРИВУ НА ОСНОВІ МЕТОДІВ ТА МОДЕЛЕЙ ЗЛИТТЯ СУПУТНИКОВИХ ДАНИХ**

05.07.12 – Дистанційні аерокосмічні дослідження

Дисертація на здобуття наукового ступеня
кандидата технічних наук

Науковий керівник:
Шелестов Андрій Юрійович
доктор технічних наук, професор

Київ – 2016

ЗМІСТ

Перелік умовних позначень	6
Вступ.....	7
Розділ 1 Сучасний стан проблеми картографування земного покриття.....	13
1.1 Картографування земного покриття на основі супутникових даних	13
1.2 Супутникові дані	20
1.3 Автоматизовані методи класифікації земного покриття	23
1.3.1 Класифікатор на основі дерева рішень	24
1.3.2 Метод опорних векторів	25
1.3.3 Нейронні мережі	26
1.4 Вибір та аналіз методу класифікації для побудови карти земного покриття	28
1.5 Методи злиття даних в задачі картографування земного покриття	30
1.6 Сервіс-орієнтована архітектура в задачах автоматизації інформаційної технології картографування земного покриття	33
1.7 Загальна постановка задачі картографування земної поверхні для великих територій	35
Розділ 2 Методи картографування земного покриття на основі злиття даних	40
2.1 Загальна схема класифікації земного покриття для великих територій на основі даних високого розрізнення	40
2.2 Злиття даних на рівні пікселів на етапі попередньої обробки.....	43
2.3 Побудова карт земного покриття для великих територій на основі методів злиття даних	47
2.3.1 Метод формування тестової та навчальної вибірки для побудови карти земного покриття	48
2.3.2 Ансамблевий метод класифікації	49
2.3.3 Алгоритм навчання та класифікації супутникових зображень	51
2.3.4 Покращення тестової та навчальної вибірок в процесі побудови карти земного покриття	54

2.4 Злиття даних на рівні прийняття рішень для отримання результуючої карти земного покриву.....	56
2.4.1 Метод злиття даних із різним просторовим розрізненням	58
2.4.2 Методи покращення карти земного покриву на основі злиття геопросторових даних.....	60
2.5 Висновки до розділу.....	62
Розділ 3 Експериментальні дослідження розроблених методів злиття супутникових даних.....	63
3.1 Порівняння кращої мережі та ансамблю	63
3.1.1 Дані	63
3.1.2 Аналіз результатів	65
3.2 Ефективність використання часових рядів супутникових даних в задачі класифікації земної поверхні	68
3.2.1 Дані	68
3.2.2 Аналіз результатів	71
3.3 Дослідження ефективності злиття оптичних та радарних даних в задачі класифікації сільськогосподарських культур.....	72
3.3.1 Дані	73
3.3.2 Аналіз результатів	75
3.4 Експериментальний аналіз розроблених методів злиття результатів класифікації на рівні прийняття рішень.....	78
3.4.1 Дані	78
3.4.2 Аналіз результатів	81
3.5 Експериментальне дослідження ефективності методів покращення карт класифікації за допомогою злиття векторних та растрових даних на рівні прийняття рішень	87
3.5.1 Дані	87
3.5.2 Аналіз результатів	88
3.6 Висновки до розділу.....	91
Розділ 4 Автоматизована інформаційна технологія картографування земного	

покриву на великих територіях у межах сервіс-орієнтованого підходу	93
4.1 Архітектура автоматизованої інформаційної технології картографування земного покриву	93
4.2 Структурна модель сервіс-орієнтованої системи в задачі картографування земного покриву	98
4.2.1 Постановка задачі структурно-функціонального аналізу сервіс- орієнтованих систем.....	99
4.2.2 Структурно-функціональний аналіз сервіс-орієнтованої системи з урахуванням типів вирішуваних завдань.....	100
4.3 Результати чисельного моделювання та оцінка ефективності передачі даних в сервіс-орієнтованій системі.....	103
4.4 Візуалізація отриманих результатів картографування земного покриву .	109
4.5 Оцінка репутації ресурсів в сервіс-орієнтованій системі	111
4.6 Висновки до розділу.....	113
Розділ 5 Практичні застосування розроблених методів злиття даних та автоматизованої інформаційної технології картографування земного покриву	114
5.1 Практична реалізація запропонованого методу злиття даних для території України	114
5.2 Карти змін	116
5.2.1 Карти змін земного покриву для території України.....	116
5.2.2 Карти змін земного покриву Донецької та Луганської областей.....	118
5.3 Порушення сівозміни.....	121
5.4 Аналіз збитків від посухи на основі злиття даних.....	123
5.4.1 Використані дані.....	124
5.4.2 Визначення пошкоджень сільськогосподарських культур на основі злиття даних	124
5.4.3 Оцінка збитків по картах класифікації.....	126
5.4.4 Практична реалізація та експериментальна перевірка запропонованого підходу оцінки збитків від посухи.....	126

5.5 Висновки до розділу.....	129
Висновки	131
References.....	134
Додатки.....	160

ПЕРЕЛІК УМОВНИХ ПОЗНАЧЕНЬ

ДЗЗ – дистанційне зондування Землі;

КА – космічний апарат;

КС – космічна система;

SVM (Support Vector Machines) – метод опорних векторів;

ККС – карти Кохонена, що самоорганізуються;

СОС – сервіс-орієнтована система;

ФЕ – функціональний елемент;

ТВ – точність виробника;

ТК – точність користувача;

СФА – структурно-функціональний аналіз;

MLP (Multilayer Perceptron) – багатошаровий персептрон;

SAR (synthetic aperture radar) – радар із синтезованою апертурою;

NNE – ансамбль нейронних мереж;

IaaS (Infrastructure as a Service) – інфраструктура як сервіс;

PaaS (Platform as a Service) – платформа як сервіс;

SaaS (Software as a service) – програмне забезпечення як сервіс.

ВСТУП

Актуальність роботи. З 2014 р. з запуском Landsat-8 та програми Copernicus Європейського космічного агентства у вільному доступі з'явилося багато різнорідних (радарних та оптичних) супутникових даних з просторовим розрізненням 10–30 м. Ці дані покривають кожну точку земної поверхні з періодичністю 5-16 днів, а значить, дозволяють розв'язувати нові класи задач моніторингу земної поверхні і будувати карти високого розрізнення.

Однією з важливих задач, які стало можливим розв'язувати на основі цих даних, є побудова карт земного покриття і землекористування розрізненням не нижче 30 м для великих територій (адміністративних областей або цілої країни) [1, 2]. До 2014 р. подібні карти земного покриття будувалися лише на основі даних середнього (250-500 м) і низького розрізнення, включали лише до 6 основних класів земного покриття і не дозволяли виділяти окремі сільськогосподарські культури. На сьогодні відома лише єдина глобальна карта з розрізненням 30 м, яка має досить низьку точність. Водночас існуючі методи обробки даних високого розрізнення ґрунтуються на обробці окремих знімків, а не часових рядів, та не передбачають спільне використання оптичних і радарних даних. Крім того, враховуючи невелику площу покриття земної поверхні одним знімком, ці методи не дозволяють будувати карти для великих територій і потребують розробки спеціальних методів геопросторового об'єднання карт, що перетинаються. Враховуючи значні об'єми супутникових даних, для розв'язання задачі картографування великих територій необхідно створювати автоматизовані технології [3, 4].

Для подолання зазначених труднощів і автоматизації процесу обробки супутникових даних необхідно розвивати методи злиття (data fusion) [5] різночасових гетерогенних даних. Саме ця задача розв'язується в дисертаційній роботі. Тому дана дисертаційна робота, присвячена розробці методів та моделей обробки та злиття супутникових даних великого об'єму, а також їх використанню для побудови карт земного покриття шляхом розробки

автоматизованої інформаційної технології картографування земного покриття для території України, є актуальною.

Зв'язок роботи з науковими програмами, планами, темами. Робота виконана у відділі космічних інформаційних технологій та систем Інституту космічних досліджень НАН України та ДКА України в межах наступних наукових тем і проектів:

- Конкурсного проекту Національної академії наук України «Розроблення методів та моделей візуалізації гетерогенної інформації для задач супутникового моніторингу надзвичайних ситуацій» (2012 р.);
- Конкурсного проекту Державного космічного агентства України «Системний проект української складової Міжнародної аерокосмічної системи глобального моніторингу» (2013 р.);
- Проекту цільової комплексної програми НАН України з наукових космічних досліджень «Розроблення науково-методичного забезпечення поліпшення точності геоприв'язки даних космічних апаратів» (2013 р.);
- Міжнародного проекту Європейського космічного агентства «Sentinel 2 for Agriculture» (2014-2017 рр.);
- Державного контракту ДКАУ «Створення інформаційної системи забезпечення користувачів аерокосмічними даними» (0108U009605) (2012 р.);
- Міжнародного проекту GEO по створенню міжнародної мережі полігонів сільськогосподарського призначення JECAM (Joint Experiment of Crop Assessment and Monitoring);
- Гранту Єврокомісії в межах Рамкової Програми FP7 «Стимулювання інновацій для глобального моніторингу сільського господарства та оцінки його впливу на навколишнє середовище як підтримка ініціативи GEOGLAM (SIGMA)».

Мета та задачі дослідження. Метою роботи дисертаційного дослідження є створення автоматизованої інформаційної технології картографування земного покриття на основі злиття супутникових та інших геопросторових даних.

Для досягнення поставленої мети необхідно розв'язати такі **задачі**:

- проаналізувати проблеми побудови карт земного покриття, існуючі методи та їх недоліки, запропонувати шляхи їх вирішення;
- вдосконалити метод класифікації земного покриття за супутниковими даними великих обсягів на основі використання ансамблю нейронних мереж та забезпечити його ефективність на основі часових рядів супутникових даних великих обсягів;
- розробити методи злиття геопросторових супутникових даних на рівні пікселів та прийняття рішень для побудови карти класифікації для великої території;
- розробити інформаційну технологію побудови карт земного покриття на основі сервіс-орієнтованого підходу.

Об'єкт дослідження: методи тематичної обробки супутникових даних, що здійснюються для отримання карти класифікації земного покриття.

Предмет дослідження: процеси та інформаційні технології обробки великих об'ємів даних спостереження Землі для великих територій.

Методи дослідження: структурно-функціональний аналіз, статистичний аналіз, геопросторовий аналіз, регресійне моделювання, інтелектуальні методи обробки даних, нейронні мережі.

Наукова новизна результатів, отриманих автором, полягає у наступному:

- **Отримав подальший розвиток** метод злиття гетерогенних супутникових даних на рівні пікселів, у якому розширюється простір ознак шляхом інтеграції різночасових даних різної природи (оптичних та радарних) для картографування земного покриття.
- **Удосконалено** метод класифікації земного покриття за супутниковими даними великих обсягів на основі використання ансамблю нейронних мереж, що визначають апостеріорну ймовірність належності кожного пікселя до відповідного класу.
- **Вперше** розроблено методи злиття даних на рівні прийняття рішень для побудови карти з використанням апостеріорної ймовірності належності

кожного пікселя до відповідного класу та маски хмарності для великої території.

- **Вперше** в межах сервіс-орієнтованого підходу розроблено автоматизовану інформаційну технологію картографування земного покриву для всієї території України.

Обґрунтованість і достовірність наукових результатів забезпечується коректним застосуванням математичного апарату, методів регресійного аналізу, використанням нейромережових моделей, а також визначенням точності отриманих карт класифікації, що оцінювались на незалежних тестових вибірках, а також шляхом порівняння з даними офіційної статистики.

Практичне значення отриманих результатів полягає в наступному:

Розроблені методи побудови карт класифікації земного покриву для великих територій реалізовано у формі інформаційної технології в межах сервіс-орієнтованої системи для автоматизованої пакетної обробки великих об'ємів супутникових даних. Зокрема ці методи використано в рамках проекту Єврокомісії SIGMA для побудови карт земного покриву на основі даних часових рядів оптичних та радіолокаційних зображень з метою подальшої побудови карт класифікації поверхні Землі. Створені карти включено до складу глобального геопросторового шару оброблюваних земель. Розроблену методологію покладено в основу методу нейромережової класифікації для проекту Sen2Agri. Результати роботи впроваджені у відділі агрометеорології Українського гідрометеорологічного центру. Також результати роботи використовуються в навчальному процесі фізико-технічного інституту НТУУ «КПІ імені Ігоря Сікорського» та в межах виконання наукової теми в Національному університеті біоресурсів і природокористування України.

Особистий внесок здобувача. Всі основні результати дисертаційної роботи отримані автором особисто. Серед 12 публікацій у фахових виданнях 1 стаття написана здобувачем одноосібно [6]. В публікаціях, написаних у співавторстві, здобувачем належать такі результати. В роботі [7] здобувачем запропоновано використовувати технології злиття супутникових даних різного

просторового розрізнення, векторні дані про межі полів та дані наземних досліджень для оцінки збитку від посухи. У роботі [8] здобувачем проведено дослідження ефективності використання ансамблю нейронних мереж. У статті [9] запропоновано метод злиття даних на основі маски хмарності та точності класифікації для побудови карт земного покриву великих територій. У роботі [10] здобувачем запропоновано метод побудови навчальної та тестової вибірки для великих територій за ретроспективними даними. У публікації [11] здобувачем запропоновано використовувати маску озимої пшениці, побудованої на основі методів злиття даних для прогнозування врожайності на рівні регіону. У роботі [12] здобувачем досліджено ефективність використання часових рядів супутникових даних. У роботі [13] з використанням методів злиття даних на рівні пікселів, що були розроблені здобувачем, побудовані карти порушень сівозміни. В публікаціях [14, 15] розроблені методи злиття даних на рівні пікселів та прийняття рішень використані для побудови карти змін. У роботах [16, 17] здобувачем запропоновано архітектуру сервіс-орієнтованої системи для підтримки прийняття рішень при надзвичайних ситуаціях.

Апробація результатів дисертації. Результати та основні положення роботи доповідалися та обговорювалися на: V міжнародній науково-практичній конференції студентів, аспірантів та молодих вчених «Інформаційні технології: економіка, техніка, освіта» (Україна, Київ, 13.11.2014) [18-21], XV міжнародній молодіжній науково-практичній конференції «Человек и космос» (Україна, Дніпропетровськ, 10-12.04.2014) [22, 23], II міжнародній науково-технічній конференції «Обчислювальний інтелект» (Україна, Черкаси, 14-18.05.2013) [24], міжнародній науково-практичній конференції «Глобальні та регіональні проблеми інформатизації в суспільстві та природокористуванні '2013» (Україна, Київ, 13-14.06.2013) [25-27], 13-тій українській конференції з космічних досліджень (Україна, Євпаторія, 02-06.09.2013) [28-30], 11-тій всеросійській відкритій конференції «Современные проблемы дистанционного зондирования Земли из космоса» (Росія, Москва ІКД РАН 11–15.11.2013 р.)

[31], четвертій міжнародній конференції «GEO-UA 2014, аерокосмічні спостереження в інтересах сталого розвитку та безпеки» (Україна, Київ, 26-30.05.2014) [32-40], всеукраїнській науковій конференції «Сучасні проблеми математичного моделювання та обчислювальних методів», (22-23.02.2013, м. Рівне, РДГУ–НУВГП) [41], 14-тій українській конференція з космічних досліджень, (8-12.09.2014, м. Ужгород) [42], II міжнародній науково-практичній конференції «Глобальні та регіональні проблеми інформатизації в суспільстві та природокористуванні 2014», (м. Київ, НУБіП України, 26-27.06.2014) [43-45], міжнародній науковій конференції «Сучасні проблеми математичного моделювання та обчислювальних методів» (м. Рівне, 19-22.02.2015) [46], міжнародній науковій конференції «Астрономічна школа молодих вчених. Актуальні проблеми астрономії і космонавтики» (м. Житомир, 20-22.05.2015) [47, 48], III міжнародній науково-практичній конференції «Глобальні та регіональні проблеми інформатизації в суспільстві та природокористуванні 2015», (м. Київ, НУБіП України, 25-26.06.2015) [49-52], International Geoscience and Remote Sensing Symposium 2015 (IGARSS 2015) [53], XIV Всеукраїнській науково-практичній конференції студентів, аспірантів та молодих вчених «Теоретичні і прикладні проблеми фізики, математики та інформатики» (м. Київ, Україна, 26-28.05.2016) [54], Materials of reports of the 5th International Conference «GEO-UA 2016» (Kyiv, 10-14.10.2016) [55].

Публікації. За матеріалами дисертації опубліковано 28 робіт, з яких 12 – це статті у журналах і збірниках наукових праць, що входять до переліку фахових видань затверджених МОН України за спеціальністю дисертації (1 з них без співавторів, а 4 включені до міжнародної наукометричної бази SCOPUS), та 16 – публікацій у матеріалах конференцій (у тому числі, міжнародних).

Структура та обсяг дисертації. Дисертація складається із вступу, п'яти розділів, висновків, списку використаних джерел. Робота містить 159 сторінок, у тому числі: 133 сторінок основного тексту, 58 рисунків, 20 таблиць, список використаних джерел із 207 найменувань на 26 сторінках.

РОЗДІЛ 1

СУЧАСНИЙ СТАН ПРОБЛЕМИ КАРТОГРАФУВАННЯ ЗЕМНОГО ПОКРИВУ

1.1 Картографування земного покриву на основі супутникових даних

Карти земного покриву (land cover maps) [56] відіграють важливу роль для осмислення динаміки екосистем та розв’язання багатьох задач супутникового моніторингу. Карта земного покриву – геопросторова інформація в растровому форматі, яка класифікує земну поверхню на обмежену кількість класів [57]. Зокрема, такі карти необхідні для визначення змін та тенденцій у землекористуванні, підвищення точності класифікації та оцінки площ, аналізу кліматичних змін та їхнього впливу на біосферу тощо.

Картою називається побудоване в картографічній проекції, зменшене, узагальнене зображення поверхні Землі або її частини, поверхні іншого небесного тіла або неземного простору, яке показує розташовані на них об’єкти у певній системі умовних знаків [58].

Для побудови карт виконується класифікація супутникових даних – це тематична обробка, яка дозволяє здійснювати автоматизоване розбиття знімків на однорідні по заданому критерію області (класи об’єктів). Клас – група об’єктів, які характеризуються низкою спільних властивостей. В даному випадку під класом будемо розуміти тип земної поверхні в даній точці, під групою об’єктів – пікселі, що відносяться до певного класу.

В процесі такого розбиття кожен піксель цифрового знімка відносять на підставі деяких статистичних критеріїв до одного з класів просторових об’єктів. Отримані при цьому зображення називаються тематичними картами [59].

Міжнародні системи класифікації земних ресурсів, які використовуються в європейських країнах, включають у себе «Стандартну статистичну класифікацію землекористування» та ряд класифікацій, які були розроблені

«Продовольчою і сільськогосподарською організацією Об'єднаних Націй» (з англ. мови FAO - Food and Agriculture Organization of the United Nations). Прикладами програм, в яких класифікація сільськогосподарського землекористування ведеться статистичним підрозділом, є:

- «Всесвітня програма перепису сільського господарства» (англ. World Programme for the Census of Agriculture) [60];
- «Глобальна програма оцінки лісових ресурсів» (англ. Global Forest Resources Assessment Programme)[61].

До програм, які були створені Європейським союзом для оптимізації існуючих систем класифікації земних ресурсів, належать:

- «Програма з аналізу та оптимізації землекористування і земельного покриття» (англ. Coordinating Land Use and Cover Data and Analyses in Europe (CLAUDE) [62];
- «Розробка гармонізованої мережевої земельної інформаційної системи на основі даних ДЗЗ» (англ. Development of a Harmonised Framework for Multi-purpose Land Cover/Land Use Information Systems Derived from Earth Observation Data (LANES) [63].

Ці програми забезпечують можливості залучення статистичних даних та даних ДЗЗ в процес класифікації земних ресурсів. Ці програми вказують на те, що навіть часткова сумісність різних систем класифікації земних ресурсів між собою є важливою [64]. Існують такі системи класифікації:

1) LCCS (англ. Land Cover Classification System) – система класифікації земного покриття [65].

Метою системи LCCS, що була винайдена в середині 90-х рр. XX ст. є стандартизація класів земних ресурсів по всьому світу. В залежності від особливостей земних ресурсів для певної території ця система може змінюватися, тому її ще називають модульною. В процесі класифікації можуть використовуватись як дані статичних спостережень, так і дані ДЗЗ [65]. У проекті AFRICOVER, що спрямований на картографування земних ресурсів Африки, система класифікації LCCS знайшла своє практичне застосування [66].

2) CLUSTERS (англ. Classification for land use statistics: EUROSTAT remote sensing programme) – класифікація Євростату для обліку землекористування на основі даних ДЗЗ [67].

Система CLUSTERS була розроблена в 1992-93 роках Євростатом та центром економічного розвитку. Ця система класифікації включає 6 класів земельного покриття на першому рівні ієрархії, які деталізуються до 66 класів на четвертому ієрархічному рівні, та орієнтована на використання даних ДЗЗ [67].

3) LUCAS (англ. Land Use/Cover Area Frame Statistical Survey) – статистичний огляд землекористування/земельного покриття [68].

Систему LUCAS було впроваджено Генеральним директором Європейської Комісії «Сільське господарство та розвиток сільської місцевості» і Євростатом в 2000-му році [68]. В рамках цієї системи виконуються наземні дослідження із залученням статистичних методів та даних ДЗЗ, отримується інформація про землекористування та земний покрив.

Дослідження LUCAS дозволяє здійснювати моніторинг змін земних ресурсів Європи за рахунок того, що проводиться на регулярній основі з часовим інтервалом у 2-3 роки. Також з кожним роком зростає кількість країн-учасників та полігонів у вибірці. У 2006 році в дослідженні брали участь 11 країн, у вибірці було 169 333 полігонів. У 2008-2009 році було залучено 25 країн, а кількість полігонів становила 263 780 [69].

Система LUCAS для земельного покриття та землекористування використовує різні системи класифікації, тобто є роздільною. Це підвищує загальну ефективність застосування даних ДЗЗ та матеріалів наземних досліджень.

4) CORINE (англ. Coordination of Information on the Environment) – координація інформації про навколишнє середовище [70].

Карти земного покриття на основі супутникових даних почали будувати у 1980-х рр. У 1985 році в Європейському Союзі розпочата програма CORINE (coordination of information on the environment) з метою збору та обміну інформацією про навколишнє середовище в межах європейських країн [71].

В рамках програми CORINE за 5 років (з 1985 по 1990 рр.) була побудована просторова база даних земельного покриття для дванадцяти країн Західної Європи, яка була названа CLC90 (CORINE Land Cover 90). Площа, охоплена базою даних CLC90, становить 2,3 млн. км², а вибірка містить 700 000 полігонів. Точність отриманої карти класифікації земного покриття становить не менше 85%. Основним джерелом даних були супутники Landsat та SPOT. Перші два етапи включають підготовку супутникових даних та додаткові геопросторові дані, на третьому та четвертому етапах відбувається побудова та постобробка отриманих карт земного покриття за допомогою середовища ГІС ArcInfo. Останній етап включав у себе перевірку результатів, їх корекцію за необхідністю [71].

В межах програми CORINE окрім побудови карт земного покриття в 1999-2000-му роках здійснювалась оцінка змін, що відбулись протягом десятиліття (проект I&CLC2000) [72]. Цим проектом управляло Європейське агентство з навколишнього середовища (EEA) та Центр загальних досліджень (JRC). Компонентами проекту I&CLC2000 були: проект IMAGE2000 (отримання даних ДЗЗ, їх ортокорекція, побудова національних і загальноєвропейських мозаїк супутникових знімків) та проект CLC2000 (визначення та інтерпретація всіх змін у земельному покритті та землекористуванні за останнє десятиліття).

В проектах CLC90 та CLC2000 використовувалась система класифікації земних ресурсів, що була розроблена в рамках проекту CORINE. Ця система містить в загальному 44 класи (табл. 1.1) [70].

Для побудови карти земного покриття традиційно використовують супутникові дані на основі знімків середнього розрізнення, зокрема MODIS або SPOT VEGETATION [73]. Такі карти використовуються для отримання глобального стану земель сільськогосподарського призначення.

В 1992 році дослідниками геологічної служби США та об'єднаного дослідницького центру Єврокомісії, була представлена глобальна карта земного покриття із просторовим розрізненням в 1 км. Формальна оцінка точності даного продукту завершена в 1998 р. [75].

Класи для карти земного покриття CORINE

№	Загальна назва класу	Назви всіх класів
1	штучні поверхні	суцільна міська забудова
		не суцільна міська забудова
		промислові та комерційні будівлі
		дороги, залізниці та прилеглі до них території
		порти
		аеропорти
		території призначені для видобутку корисних копалин
		звалища
		будівельні майданчики
		парки та сквери
		території призначені для спорту та відпочинку
2	сільськогосподарські землі	не зрошувані орні землі
		зрошувані орні землі
		рисові поля
		виноградники
		сади та плантації ягід
		насадження олив
		пасовища
		однорічні культури з вкрапленнями багаторічних
		складні зразки культивування
		сільськогосподарські землі зі значними територіями вкритими природною рослинністю
3	ліси та напівприродні території	сільськогосподарські землі частково вкриті лісом
		широколисті ліси
		хвойні ліси
		мішані ліси
		природні пасовища
		мохові болота і пустки
		склерофільна рослинність
		чагарники, що переходять у ліс
		пляжі, дюни та піски
		голі скелі
		землі з незначною рослинністю
4	болота	гарі
		льодовики та вічні сніги
		болота
		торфовища
		солончаки
5	водні об'єкти	солонці
		припливні рівнини
		річки, канали
		озера, ставки
		лагуни
		естуарії
		моря та океани

Побудовані карти CORINE є у вільному доступі на сайті «Copernicus Land Monitoring Services» (рис. 1.1) [74].

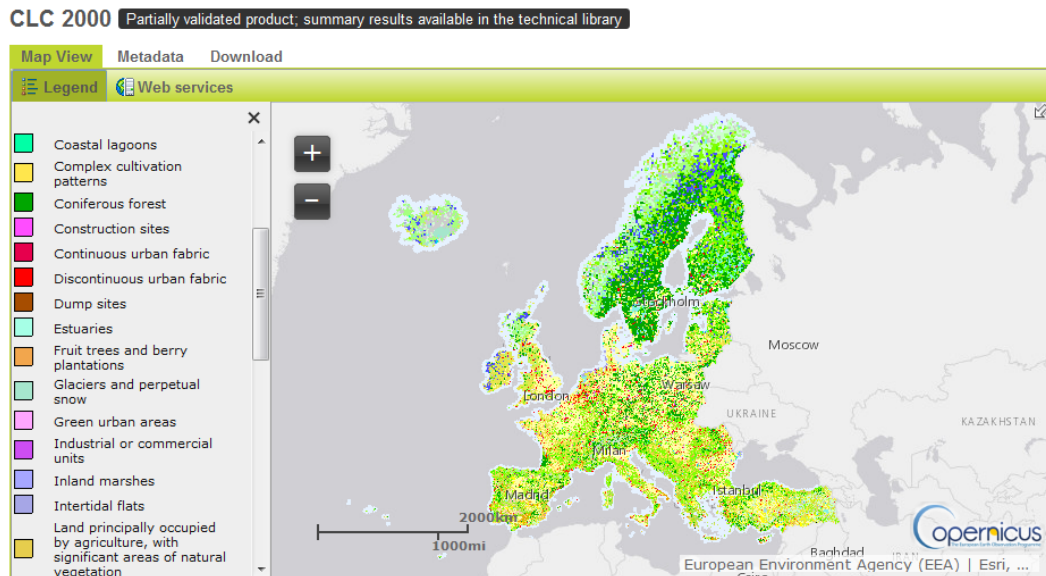


Рис. 1.1 Карта земного покриття CORINE (2000 рік)

Проект GlobCover 2009 є другим 300-м глобальним продуктом, створеним автоматизованою класифікацією на основі MERIS FR часових рядів (табл. 1.2). В основі обрахунків використовується 22 класи земного покриття, що зареєстровані і використовуються у всьому світі [76]. Основними недоліками існуючих продуктів є те, що вони містять всі землі сільськогосподарського призначення, а не оброблювані землі для поточного року, а також мають низьке просторове розрізнення та малу точність побудованих карт земного покриття.

З наявністю даних Landsat-8 і Sentinel та можливістю їх сумісного використання [77], виникає можливість будувати карти земного покриття із високим просторовим розрізненням. Однак, більшість досліджень з картографування земного покриття, використовуючи дані високого розрізнення (наприклад, Landsat 5/7, SPOT, AWiFS), були проведені на регіональному рівні [78 – 80].

Побудова карт на основі даних низького розрізнення спричиняє недооцінку або переоцінку площ певних типів земного покриття чи їхніх змін. Тому, з появою у вільному доступі часових рядів даних супутника Landsat надзвичайно актуальною стала задача побудови глобальних і регіональних карт земного покриття на основі супутникових зображень високого (30 м) просторового розрізнення. В 2013—2014 рр. з'явились кілька таких карт [81,

82], які в роботі USGS [56] названі картами «нового покоління».

Як видно із таблиці 1.2, останні три продукти отримані на основі даних високого розрізнення (наприклад, для побудови GlobeLand30-2010 використані супутникові дані Landsat, що покривають всю Землю з просторовим розрізненням 30 м) [56]. Для побудови карти земного покриття розроблено автоматизований підхід, який базується на злитті попиксельних та об'єктно-орієнтованих методів класифікації (ПОК-based). Продуктивність заснованого підходу (ПОК-based), була оцінена з використанням восьми обраних ділянок з різними покриттями на п'яти різних континентах. Була досягнута загальна точність класифікації більше 80% [82]. Проте, одним із основних недоліків побудованої карти є те, що серед отриманих класів немає відокремлених земель сільськогосподарського призначення, що використовувалися у поточному році. Також даний підхід не враховує територіальних відмінностей, пов'язаних із кліматом та різномірністю посівів, що проявляється для території України (наприклад пустеля в Херсонській області віднесена до земель сільськогосподарського призначення).

Національна служба сільськогосподарської статистики (NASS) Міністерства сільського господарства США (USDA) створює карту земного покриття (Cropland Data Layer - CDL) у растровому форматі, що має географічну прив'язку та специфікацію рослинності [83]. Продукт CDL містить карти для 47 держав із просторовим розрізненням супутникових знімків 56 м. Також було додано зусиль, щоб створити глобальний продукт із оброблюваними землями, на основі даних Landsat TM і ETM+ [84, 85]. Для цього автори побудували глобальний ґрунтово-рослинний продукт із просторовим розрізненням 30 м під назвою OT-GLC (Fine Resolution Observation and Monitoring of Global Land Cover). Точність виробника (ТВ) і точність користувача (ТК) для класу оброблюваних земель складає 75,25% та 55,62%, що не забезпечує надійного використання [86]. Хоча продукти GLC використовуються, їх якість далека від задовільної для багатьох застосувань [87 – 89].

Карти глобального земного покриття

Джерело даних	Рік	Розрізнення	Посилання
Mathews Global Vegetation/Land Use	1983	1 км x 1 км	http://www.giss.nasa.gov/
Olson Land Cover and Vegetation	1983	0.5 км x 0.5 км	http://www.grid.unep.ch/
Willson and Henderson-Sellers Global Land Cover	1985	1 км x 1 км	http://www.ngdc.noaa.gov/
DeFries/Townshend-Global Land Cover	1995	10 км x 10 км	http://glcf.umiacs.umd.edu/
GLCC (IGBP DISCover)	1997	1 км x 1 км	http://edcdaac.usgs.gov/
UMD Land Cover	2000	1 км x 1 км	http://glcf.umiacs.umd.edu/
MODIS Land Cover	2003	1 км x 1 км	http://edcdaac.usgs.gov/
Vegetation Continuous Fields	2003	1 км x 1 км	http://glcf.umiacs.umd.edu/
GLC-2000	2003	1 км x 1 км	http://www.gvm.sai.jrc.it/
Vegetation Continuous Fields	2003	500 м x 500 м	http://glcf.umiacs.umd.edu/
MODIS Land Cover	2008	500 м x 500 м	http://edcdaac.usgs.gov/
GLOBCOVER	2009	300 м x 300 м	http://www.esa.int/
GeoCover LCTM	2000	30 м x 30 м	http://www.earthsat.com/
China Global Land Cover	2012	30 м x 30 м	http://www.globallandcover.com/
GEO US Global Land Cover	2013	30 м x 30 м	http://landcover.usgs.gov/

Отже, навіть за умови валідації цих глобальних карт на численних географічно розподілених полігонах їхня точність для конкретних регіонів є недостатньо високою.

1.2 Супутникові дані

Зображення дистанційного зондування з космосу завжди були перспективним джерелом інформації для отримання карти земного покриття. Головним чином це пояснюється можливістю своєчасно отримувати

зображення і забезпечувати повторюваність вимірювань для великих територій.

За кілька останніх років стрімко зростає кількість новітніх інструментів для отримання, збереження та обміну інформацією про навколишнє середовище – запущені супутники із роздільною здатністю 10-300 м (Landsat-8, PROBA-V, Sentinel-1 та Sentinel-2), що призвело до накопичення великих об'ємів даних різної природи, які є цінним джерелом інформації для дослідження Землі. Оскільки в останні роки актуальною стала задача побудови карт земного покриття для великих територій з використанням даних високого просторового розрізнення, то це стало можливим завдяки появі часового ряду даних із нових супутникових апаратів.

Серед великої кількості сучасних супутників, що використовуються для картографування земного покриття, є супутники із високим просторовим розрізненням, більшість із яких є лише в платному доступі. Проте за останні роки з'явилися безкоштовні супутникові дані, як радарні так і оптичні (табл. 1.3).

Таблиця 1.3 –

Основні характеристики космічних апаратів

Супутник	Розмір сцени або ширина захвату	Тип сенсора	Просторове розрізнення (м)	Ціна
Sentinel-1	250x250 км ²	Радарний	10	безкоштовно
Sentinel-2	100 км	Оптичний	10	безкоштовно
Landsat-8	185x180 км ²	Оптичний	30	безкоштовно
Proba-V	2250 км	Оптичний	100	безкоштовно

Одним із застосувань супутникових даних Landsat є побудова карт класифікації рослинного покриття. Супутник Landsat покриває поверхню Землі кожні 16 днів. Проведені дослідження підтверджують ефективність застосування Landsat-8 для класифікації земного покриття [77, 90 – 92]. Proba-V забезпечує покриття по всій планеті кожні два дні, ширина захвату однієї сцени 2250 км, що дає можливість отримувати часові ряди даних, які можуть бути використанні для побудови карт земного покриття, відстеження динаміки зміни

вегетації рослин та ін. [93]. При розв'язанні задачі картографування земного покриву нові можливості відкривають безкоштовні супутникові дані Sentinel - 1,2, оскільки покривають територію кожні 12 днів та надають дані із просторовим розрізненням 10 м [94] (рис. 1.2).

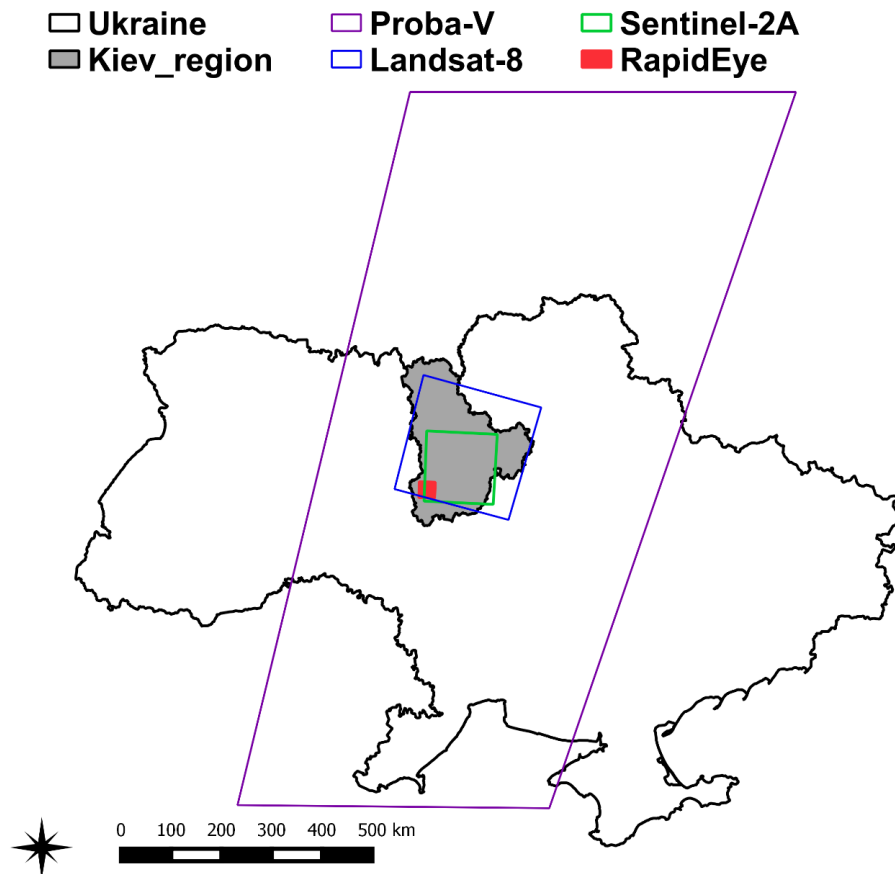


Рис. 1.2 Проблема великих об'ємів даних

У планети Земля є багато супутників, які знімають велику кількість гетерогенної інформації в різних просторових і часових розрізненнях. Наявність такої великої кількості даних, включаючи європейські та американські супутники, забезпечує появу віртуальних угруповань, але вони всі різні і тому необхідно їх гармонізувати. Тобто ці дані необхідно привести до одного виду, щоб їх можна було використовувати. Тому при побудові карт земного покриву виникає задача злиття супутникових даних для того, щоб їх сумісно використовувати і будувати карти на рівні пікселів для кожної зі сцен. Побудувати карту для великої території неможливо, оскільки кожний знімок

покриває різну територію за площею та з використанням різної кількості даних. Тому для кожної спільної частини території необхідний свій класифікатор, після чого відбувається злиття даних на рівні прийняття рішень.

Використання наведених супутникових даних дає можливість будувати карти земного покриття із високим просторовим розрізненням та із певною періодичністю. Одним із недоліків оптичних супутникових даних є те, що знімки можуть бути захмареними, що не дає можливості достовірно використовувати ці дані. Радарні дані в даному плані мають свою перевагу, так як дають можливість отримувати знімки при будь-яких погодних умовах, проте отримані дані зашумлені в порівнянні із оптичними (навіть після проведення попередньої обробки).

Тому, виникає необхідність розробки методів та моделей злиття супутникових даних для забезпечення обробки великих об'ємів даних та побудови карти земного покриття.

1.3 Автоматизовані методи класифікації земного покриття

При побудові карти земного покриття на основі даних ДЗЗ, одним із основних етапів є класифікація супутникових даних. Супутникові знімки не являють собою карти, а лише дають можливість визначити на їх основі класи земного покриття. Можливий спосіб побудови карти на основі фотоінтерпретації для невеликої кількості полігонів, але даний процес складний і довготривалий, особливо коли необхідно побудувати карту земного покриття для великої території. Тому, виникає необхідність проведення автоматизованої класифікації для отримання результуючої карти. Для цього необхідно використати метод класифікації, який забезпечить достовірні результати.

Методам класифікації земної поверхні за даними ДЗЗ присвячена велика кількість статей як у вітчизняній, так і в англomовній літературі [77, 90, 95 – 101]. Однак у перерахованих та інших публікаціях найчастіше ставиться задача класифікації із використанням одного супутникового зображення, що має

низьке просторове розрізнення (наприклад, дані MODIS [102]). При цьому задача класифікації найчастіше розв'язується для окремого району та не аналізується можливість поширення отриманих результатів на більш великі території [9].

У джерелах, де використовуються методи побудови карт земного покриття, надаються переваги наступним методам класифікації: максимальної правдоподібності [103], найближчого сусіда [104], логістичної регресії [105], методам, що базуються на правилах прийняття рішень [78], деревам прийняття рішень [106, 107], нейромережевому методу [108], методу опорних векторів [109], спектральним методам розшарування [110] та біологічним алгоритмам [111], які базуються на основі попиксельної або на об'єктно-орієнтованій класифікації [107, 112]. Деякі методи залишилися на рівні розробок [113], в інших [114], [115], [116], [117] була встановлена важливість виділяти багаторічні культури [79].

1.3.1 Класифікатор на основі дерева рішень

На сьогодні одним з популярних підходів до розв'язання задач класифікації є дерева рішень [33, 118]. Саме цей підхід успішно використовується в США для класифікації територій земного покриття за супутниковими даними [119].

Дерева рішень зазвичай використовуються для розв'язку задач класифікації даних. Випадок, в якому варто застосовувати дерева прийняття рішень, звичайно виглядає так: є багато випадків, кожний з яких описується деяким кінцевим набором дискретних атрибутів. В кожному з випадків є значення деякої (невідомої) булевої функції, залежної від цих атрибутів. Задача полягає у створенні досить економічної конструкції, яка б описувала цю функцію та дозволяла класифікувати нові дані, що надходять ззовні. Будь-яке дерево рішень по суті представляє собою деревоподібний граф.

Визначення класу та значення для вибірки здійснюються шляхом

переходу по дереву, починаючи від кореневого вузла. Кожен етап переходу здійснюється з урахуванням результатів перевірки визначених умов.

В останні роки дерева рішень успішно використовуються для класифікації земного покриття на основі даних дистанційного зондування [120, 121].

Перевага використання методу на основі дерев рішень полягає в тому, що потрібен малий обсяг інформації для навчання, тому він займає в пам'яті небагато місця. Після того як необхідні значення представлені у вигляді дерева, найчастіше відпадає необхідність зберігати вибірки даних, що застосовувалися для навчання дерева.

Проте слід зазначити деякі недоліки. Дерева рішень добре підходять для пакетної обробки наборів даних. Але після того, як виникла необхідність проводити їх навчання в оперативному режимі, існуючі алгоритми можуть виявитися досить громіздкими (які вимагають великого обсягу пам'яті і є складними в реалізації). Більшість алгоритмів побудови дерев рішень є «жадібними». Це означає, що якщо один раз змінна була обрана і по ній було зроблено розбиття на підмножини, то алгоритм не може повернутися назад і вибрати іншу змінну, яка дала б краще розбиття. Тому на етапі побудови не можна сказати, чи дасть обрана змінна оптимальне розбиття в кінцевому підсумку [122].

Також метод дерев рішень схильний до перенавчання. Враховуючи вказані недоліки використання методу дерев рішень для побудови карт земного покриття для великих територій є неефективним.

1.3.2 Метод опорних векторів

Метод опорних векторів (Support Vector Machines, SVM) був розроблений В. Вапніком [123]. У найпростішому випадку метод SVM полягає в знаходженні в просторі ознак гіперплощини, яка ділить його на дві частини: в одній знаходяться всі позитивні приклади, а в іншій – всі негативні приклади.

При цьому серед всіх таких гіперплощин знаходиться та, для якої мінімальна відстань до найближчих зразків є максимальною.

Метод опорних векторів розв'язує задачу класифікації за допомогою визначення найменшої відстані між межею рішень і будь-яким зразком. Межа рішень вибирається тоді, коли відстань є максимальною. Відстань визначається як перпендикуляр між межею рішень і найближчими точками даних. Максимальне значення призводить до певного вибору межі рішень.

До переваг методу SVM можна віднести такі фактори: швидкість навчання, теоретична і практична обґрунтованість методу; загальний підхід до багатьох задач; стійкість розв'язку, відсутність проблем з локальними мінімумами; не схильний до проблем перенавчання; працює з будь-якою кількістю вимірювань.

Також можна зазначити такі недоліки методу опорних векторів: невисока продуктивність в порівнянні з більш простими методами; відсутність загальних рекомендацій по підбору параметрів; проблеми з нелінійними перетвореннями; складності з інтерпретацією результату; робота з невеликою кількістю векторів.

Враховуючи перераховані переваги та недоліки методу опорних векторів, а також те, що цей метод є ефективним для задач з обмеженою вибіркою для навчання [124], можна зробити висновки, що він не є оптимальним, оскільки при розв'язанні задачі дисертаційного дослідження використовується великий обсяг геопросторових даних.

1.3.3 Нейронні мережі

На сьогоднішній день при виборі підходу для класифікації земного покриття за даними дистанційного зондування Землі все частіше віддають перевагу нейромережевим методам. Нейронні мережі забезпечують адаптивний і стійкий підхід, що дає змогу аналізувати та узагальнювати інформацію, який не вимагає апріорних знань про статистичний розподіл даних.

Штучні нейронні мережі (ШНМ) – математичні моделі, а також їх

програмні або апаратні реалізації, побудовані за принципом організації та функціонування біологічних нейронних мереж – мереж нервових клітин живого організму. Це поняття виникло при вивченні процесів, що протікають в мозку, а саме при спробі моделювання цих процесів [125, 126].

Навчання нейронних мереж – це багатопараметрична задача нелінійної оптимізації [108]. Можливість навчання – одна з головних переваг нейронних мереж перед традиційними алгоритмами. Технічно навчання полягає в знаходженні коефіцієнтів зв'язків між нейронами. В процесі навчання нейронна мережа здатна виявляти складні залежності між вхідними даними та вихідними, а також виконувати узагальнення. Це означає, що у випадку успішного навчання, мережа зможе правильно визначити вірний результат на підставі даних, які були відсутні в навчальній вибірці, а також неповних і/або частково спотворених даних.

Використання нейронних мереж є ефективним підходом для статистичного розпізнавання образів. В практичних задачах класифікації найчастіше використовуються нейронні мережі прямого поширення, відомі як багат шаровий перцептрон або MLP (Multilayer Perceptron), для навчання яких використовуються різні модифікації методу градієнтного спуску [127]. Згідно з теоремою Колмогорова, нейронна мережа MLP є універсальним класифікатором [128].

В процесі навчання багат шарового перцептрону фактично проводиться пошук значень вагових коефіцієнтів і порогів кожного елемента, які б мінімізували помилку прогнозу, який видається мережею [129].

Підходи на основі нейронних мереж мають явну перевагу над статистичними методами класифікації, тому що вони не є параметричними. До інших переваг нейронних мереж включають паралельні обчислення, здатність оцінити нелінійну залежність між вхідними даними і необхідними виходами та швидка можливість узагальнення результатів [130 – 132].

Недоліком методу нейронних мереж є малий обсяг даних. Нейронна мережа потребує навчальну вибірку великого обсягу для більш коректного

навчання.

1.4 Вибір та аналіз методу класифікації для побудови карти земного покриття

В роботі [121] проведено експеримент, метою якого було порівняти методи класифікації та обрати найкращий класифікатор для побудови карт земного покриття. Даний експеримент проведено на п'яти JESAM полігонах, розташованих в Україні, Росії, Аргентині, Китаї і Бразилії. Проведений експеримент дає можливість оцінити ефективність методів класифікації з виявленням їх сильних сторін, а також оцінити їх часті помилки із урахуванням просторового розрізнення використаних супутникових даних. Розглянемо детальніше проведений експеримент. Він включає наступні методи: аналіз та ансамблева класифікація часових рядів (TSAEC), ансамбль нейронних мереж (NNE), класифікація за допомогою дерев прийняття рішень (DT), метод картографування оброблюваних земель для великих територій (LSAM), класифікація, що базується на знаннях про оброблювані землі сільськогосподарського призначення (KBC). Для того, щоб порівняти методи, вхідні супутникові дані були однаковими для різних полігонів. Зібрані наземні дані учасниками JESAM випадковим чином розділені на навчальні (50% полігонів) та тестові (інші 50% полігонів). Полігони розділені таким чином, щоб виключити перетин між навчальним та тестовим набором даних. Для експерименту використовувалися знімки низького та високого просторового розрізнення (зокрема Landsat-8).

Розроблений метод класифікації кожної із країн був апробований на даних всіх п'яти тестових полігонів JESAM. Оцінка точності карт здійснювалася на основі матриці невідповідності (що будується на тестових незалежних даних). Для всіх тестових полігонів загальна точність ансамблю нейронних мереж (NNE - neural network ensemble) більша 90% (в діапазоні від 90% для китайського тестового полігону до 98% і 98% для українських і

російських, відповідно). Даний метод, перевершує інші підходи з точки зору загальної точності ОА для чотирьох з п'яти сайтів JECAM (рис. 1.3).

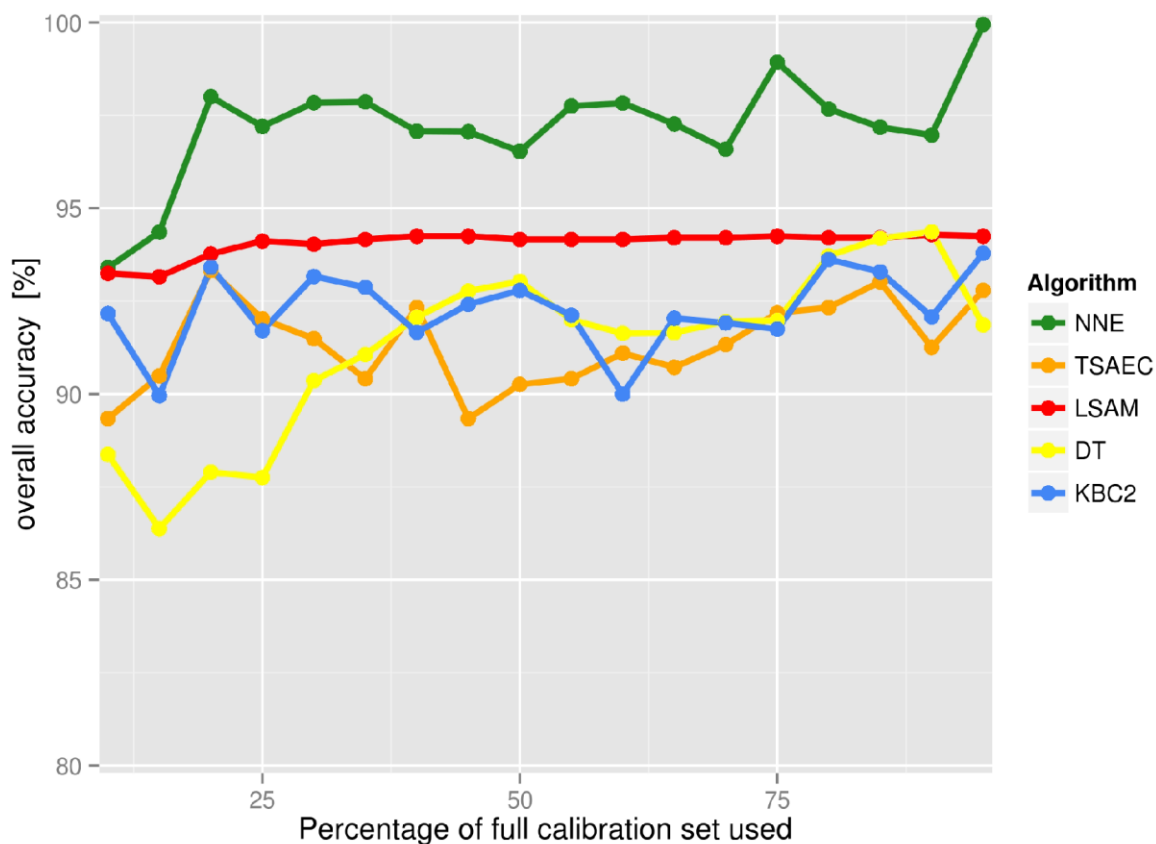


Рис. 1.3 Залежність загальної точності (ОА) для тестового набору даних (виражено у відсотках від повного розміру навчання) для JECAM полігонів

Результати експерименту показують, що NNE в більшості випадків випереджає інші досліджувані методи [4, 8, 121, 133]. Хоча на етапі навчання NNE може бути ресурсоємним і забирає багато часу, але це стає менш проблематичним із використанням високопродуктивних обчислень [16, 134 – 136].

Враховавши те, що для побудови карт земного покриття для великої території використовуємо часовий ряд супутникових даних, в наслідок чого накопичується великий обсяг даних, можна зробити висновок, що нейронні мережі є оптимальними для розв'язання поставленої дисертаційної задачі.

1.5 Методи злиття даних в задачі картографування земного покриття

За останні кілька років, у зв'язку з появою нових супутників, дані дистанційного зондування Землі стали широко доступними для користувачів. Проте, інформація, яку містить кожний окремий супутник, може бути неповною, пошкодженою, неточною для застосування тощо. В такому випадку додаткові джерела даних можуть надати інформацію, якої не вистачає. Таким чином, злиття різної інформації може надати більш точну картину для аналізу даних дистанційного зондування [137].

Перевагами злиття даних є: можливість поліпшення просторового розрізнення супутникового зображення, покращення якості зображення, підвищення точності класифікації, можливість замінювати відсутню інформацію в одному зображенні на інформацію з іншого зображення та інше. Супутникові дані відрізняються просторовим, радіометричним розрізненням та площею покриття. Беручи до уваги стрімко зростаюче різноманіття космічних апаратів, виведених на орбіту Землі протягом останніх років, виникає актуальна задача, що полягає в необхідності обробки великих об'ємів даних, використовуючи розподілені обчислювальні ресурси та високопродуктивні обчислення [138 – 141].

Зважаючи на все вищесказане, можна зазначити, що для своєчасної обробки великих об'ємів різнорідних даних та створення якісних тематичних карт необхідно розробити швидкі, надійні і точні методи класифікації та пост-обробки супутникових зображень, які ефективно справлятимуться з різнорідною геопросторовою інформацією [142 – 144].

Прив'язаність тематичних карт землекористування до конкретного часового періоду, а також їх просторова обмеженість, викликають необхідність їх періодичного оновлення, а також залучення до процесу картографування інших можливих супутникових даних.

При розв'язанні задачі картографування земної поверхні для території, яку покривають дані із декількох джерел, виникає задача злиття даних (data

fusion) для отримання результуючої карти класифікації [5].

Злиття даних (data fusion) – процес інтеграції даних і знань різної природи, що стосуються одного і того ж об'єкту. Використання даних дистанційного зондування, особливо даних декількох джерел, залишається складним завданням з багатьох причин, таких як: різне просторове розрізнення даних, супутникові дані нерівномірно покривають територію, дані можуть містити хмари та тіні від хмар, що породжує необхідність використання інших супутникових даних, неможливість покриття одним знімком великої території при побудові карт земного покриття.

Загалом, дистанційні методи злиття можуть класифікуватися за трьома рівнями: рівень пікселів, рівень ознак та рівень прийняття рішень (рис. 1.4) [5].

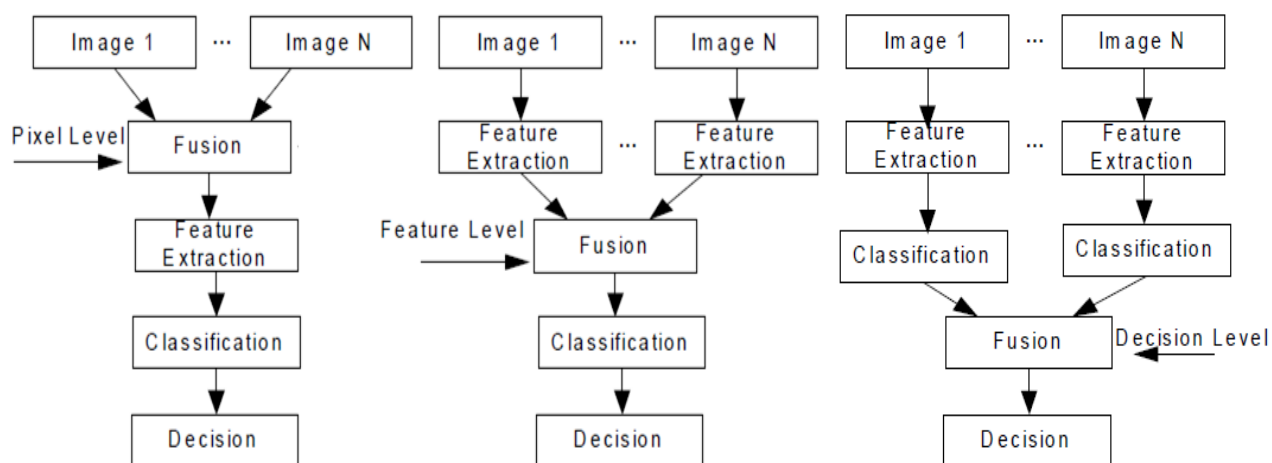


Рис. 1.4 Класифікація рівнів методів злиття даних

Злиття даних на рівні пікселів є поєднанням вихідних даних з декількох джерел в одне, яке, як очікується, буде більш інформативним, ніж будь-які із вихідних даних, отриманих в різний час.

Методи злиття на рівні ознак вимагає вилучення характерних особливостей із вхідних зображень. Характеристикою може бути інтенсивність пікселів або особливість текстури зображення. Особливості включають в себе краї, область, форму, розмір, довжину або зображення сегменту, а також особливості з аналогічними характеристиками інтенсивності в зображеннях,

щоб була можливість злиття різних типів зображень одного і того ж географічного району. Ці характеристики потім об'єднуються з аналогічними характеристиками, які присутні в інших вхідних зображеннях за допомогою заздалегідь визначеного процесу відбору для формування кінцевого зображення. Злиття на рівні ознак важко досягти, коли набори ознак отримані з різних алгоритмів і джерел даних.

Злиття даних на рівні ознак є особливо важливим, коли число доступних спектральних каналів стає настільки великим, що неможливо аналізувати кожний канал окремо. Рівень прийняття рішень об'єднує результати з декількох алгоритмів для отримання остаточного рішення (наприклад результати від різних алгоритмів або оцінки). У злитті на рівні прийняття рішень окремі зображення спочатку класифікуються окремо, використовуючи будь-який метод класифікації зображень, а потім діє правило прийняття рішень, яке усуває неясності в розподілі класів та застосовується для побудови узагальненої карти класифікації. Методи злиття даних на основі прийняття рішень включають в себе способи голосування, статистичні методи, методи нечіткої логіки та ін.

Кожен з зазначених способів злиття даних спрямований на досягнення конкретної мети, в нашому випадку – на побудову карти класифікації з використанням великих обсягів інформації. Отже, злиття даних проводиться для подальшої класифікації зображень, тому основним критерієм є точність результатів класифікації.

В контексті даного дисертаційного дослідження з розглянутих методів злиття даних буде застосовуватися метод на рівні пікселів та на рівні прийняття рішень. Метод злиття даних на рівні ознак не застосовується, тому що використовується вся наявна геопросторова інформація, а не окремі її характеристики, чи продукти, що побудовані на її основі, наприклад, NDVI (normalized difference vegetation index).

1.6 Сервіс-орієнтована архітектура в задачах автоматизації інформаційної технології картографування земного покриття

Стрімкий розвиток комп'ютерних технологій призвів до того, що для виконання поточних завдань над великими масивами даних набагато вигідніше, замість підтримки власних підрозділів з обчислювальною інфраструктурою, використовувати одну спільну інфраструктуру для виконання однотипних завдань. В основу такої інфраструктури покладений сервіс-орієнтований підхід, коли цільовий інформаційний ресурс може розглядатися як набір окремих сервісів [145, 146]. Розробка таких сервіс-орієнтованих підходів ефективна при побудові карт земного покриття для великих територій.

Сервіс-орієнтована архітектура – це парадигма, призначена для проектування, розробки та управління дискретними одиницями логіки (сервісами) в обчислювальному середовищі [146]. Застосування такого підходу вимагає від розробників проектування власного програмного забезпечення. Розробники при побудові нових компонентів системи можуть використовувати вже існуючі блоки, які будуть легко інтегруватися в будь-яку систему незалежно від деталей їх реалізації.

Сервіс – це абстрактний засіб доступу до конкретних додатків або баз даних. Призначенням сервіс-орієнтованої архітектури є надання користувачам можливості отримувати кращі в своєму класі компоненти без аналізу способу доступу до них, а також їх розташування, на основі стандартизованих інтерфейсів. В якості сервісу може виступати і інфраструктура, і платформа, і програмне забезпечення, залежно від завдань, які стоять перед користувачем [147].

Інфраструктура як сервіс (Infrastructure as a Service – IaaS) – це побудова комп'ютерної інфраструктури (як правило, підтримуючої віртуалізацію) як сервісу, доступ до якої реалізований на основі концепції хмарних обчислень. IaaS охоплює сервіси, які можна пропонувати користувачам на основі принципу погодинної оплати. Це можуть бути такі сервіси, як надання

обчислювальних і мережних ресурсів, а також ресурсів зберігання даних.

Платформа як сервіс (Platform as a Service – PaaS) – надання інтегрованої платформи для розробки, тестування, розгортання та підтримки веб-сервісів, доступ до якої реалізований на основі хмарних обчислень.

Зараз більшість додатків розробляються в одному середовищі, тестуються в іншому, а розгортаються в третьому. Використання концепції PaaS дозволяє виконувати всі ці етапи розробки програмного забезпечення в одному оточенні, що істотно спрощує весь цей процес [148].

Програмне забезпечення, як сервіс (Software as a service – SaaS) – це надання програмного забезпечення через Інтернет, коли постачальник розробляє веб-додаток і самостійно адмініструє його. Основною перевагою використання моделі SaaS для користувача є відсутність капіталовкладень, пов'язаних з установкою, оновленням і підтримкою обладнання та відповідного програмного забезпечення, а також утриманням штату технічних працівників [149].

Сервіс-орієнтована архітектура характеризується наступними принципами [150, 151]:

- сервіси, як компоненти інформаційної системи, є незалежними від платформи мови програмування, операційної системи та інших особливостей технічної реалізації [152 – 155]. Сервіси взаємодіють між собою і з іншими додатковими службами, використовуючи відкриті стандарти, наприклад, протокол SOAP (Simple Object Access Protocol);

- кожен сервіс як абстрактне уявлення функціональності реалізує окрему функцію, яка є логічно-відокремленою і виконує автономні завдання без використання інших сервісів або програмних компонентів;

- низька зв'язність сервісів з іншими компонентами системи.

Таким чином, сервіси є автономними, незалежними від платформи об'єктами, пов'язаними між собою вузькими, чітко визначеними інтерфейсами.

У загальному випадку сервіс-орієнтована архітектура формується користувачами, постачальниками сервісів і брокером сервісів (рис. 1.5) [156].

Постачальник сервісу публікує інформацію про представлені ресурси у брокера сервісів. У свою чергу за допомогою звернення до брокера користувачі можуть вибрати і використовувати необхідний сервіс. При цьому умови надання сервісу визначаються його постачальником у відповідному контракті [156].

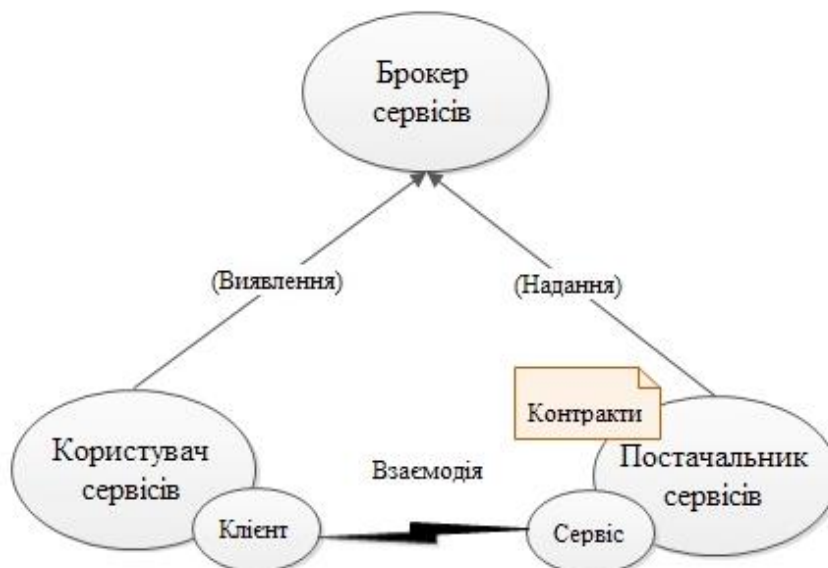


Рис. 1.5 Принципи взаємодії в межах сервіс-орієнтованої архітектури

Однак, незважаючи на перехід від традиційних систем до сервіс-орієнтованих, вимоги до забезпечення якості наданих сервісів залишаються незмінними. Такі системи повинні бути надійними, безпечними, доступними й ефективними у використанні [157]. На жаль, домогтися виконання всіх цих вимог для сервіс-орієнтованих систем складніше, ніж для традиційних систем, оскільки сервіси можуть надаватися сторонніми розробниками. Саме для цього необхідно виявити вузькі місця, які призводять до зниження ефективності роботи сервіс-орієнтованої системи і продуктивності роботи системи в цілому.

1.7 Загальна постановка задачі картографування земної поверхні для великих територій

Таким чином, поява даних високого розрізнення відкриває великі можливості та водночас призводить до виникнення нових проблем. Зокрема, будувати карти окремих культур для великих територій, що раніше не представлялося можливим, за рахунок використання часового ряду

супутникових даних. В той же час такі технології побудови карти земного покриття створені лише для невеликих територій (що покриваються одним знімком високого розрізнення), а для великих така технологія відсутня, тому виникає необхідність в її розробці. Відсутність такої технології спричиняється наступними проблемами (вирішення яких пропонується в даному дисертаційному дослідженні):

- неможливість покриття одним знімком високого розрізнення великої території (рис. 1.6-1.8);
- нерівномірне покриття супутниковими зображеннями цільової території, що призводить до необхідності використання різних ознак для різних областей (регіонів). Так, на рис. 1.6-1.8 видно, що різні території України в різні роки були зафіксовані супутником різну кількість разів. Для того, щоб покрити всю територію України та використовувати часовий ряд даних, необхідно об'єднувати дані різних супутникових апаратів;

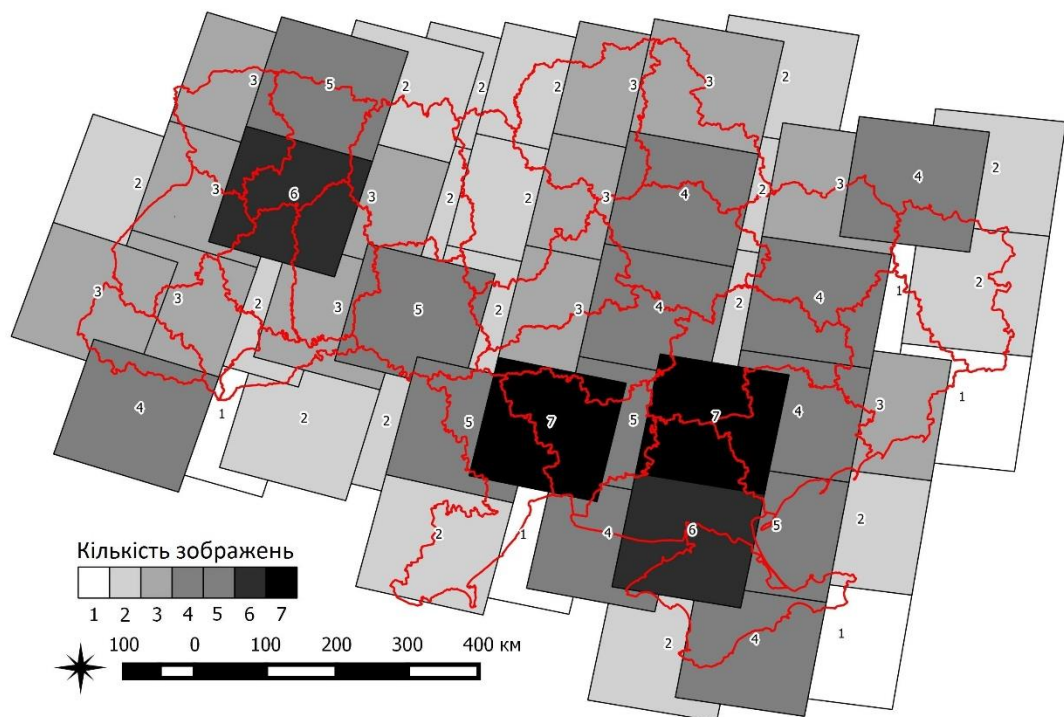


Рис. 1.6 Покриття території України сценами Landsat-5/7, що використовуються для класифікації на основі часового ряду 2010 р.

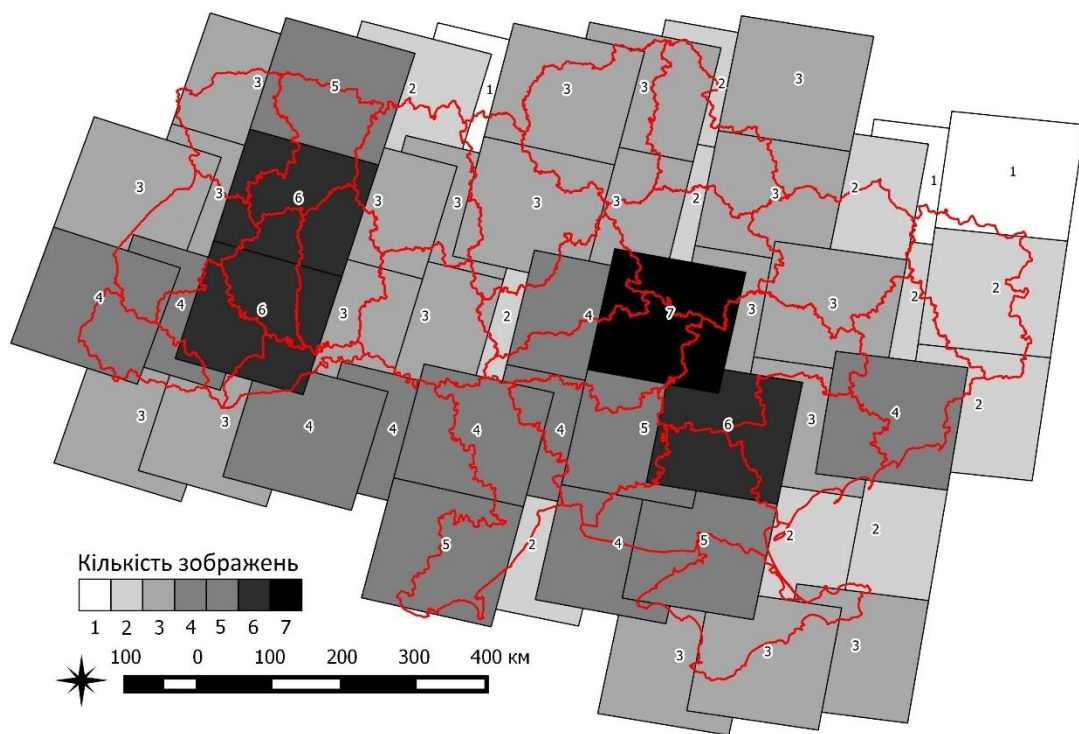


Рис. 1.7 Покриття території України сценами Landsat-5/7, що використовуються для класифікації на основі часового ряду 2000 р.

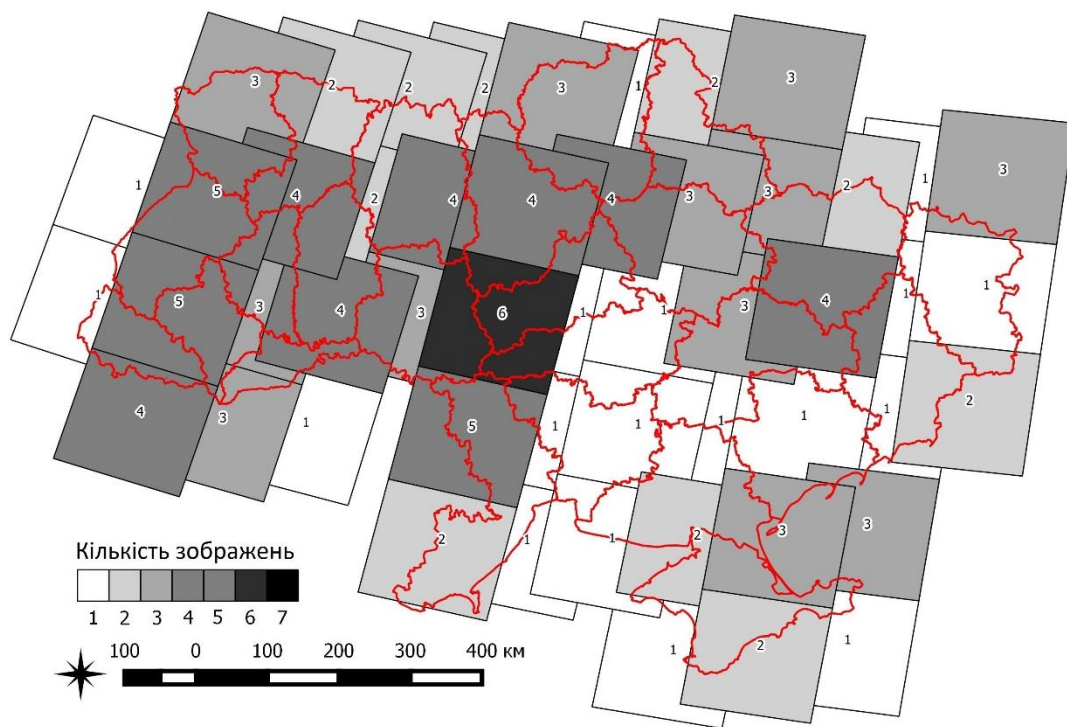


Рис. 1.8 Покриття території України сценами Landsat-4/5, що використовуються для класифікації на основі часового ряду 1990 р.

- необхідність урахування локальних характеристик досліджуваного

регіону (наприклад, інтенсивність і фази розвитку сільськогосподарських культур можуть сильно варіюватися залежно від агрокліматичних зон). Наприклад, для території України період вегетації між однаковими сільськогосподарськими культурами відрізняється, на півдні вегетація починається раніше ніж на півночі;

- наявність пропусків даних через хмарність та тіні на оптичних зображеннях (рис. 1.9);

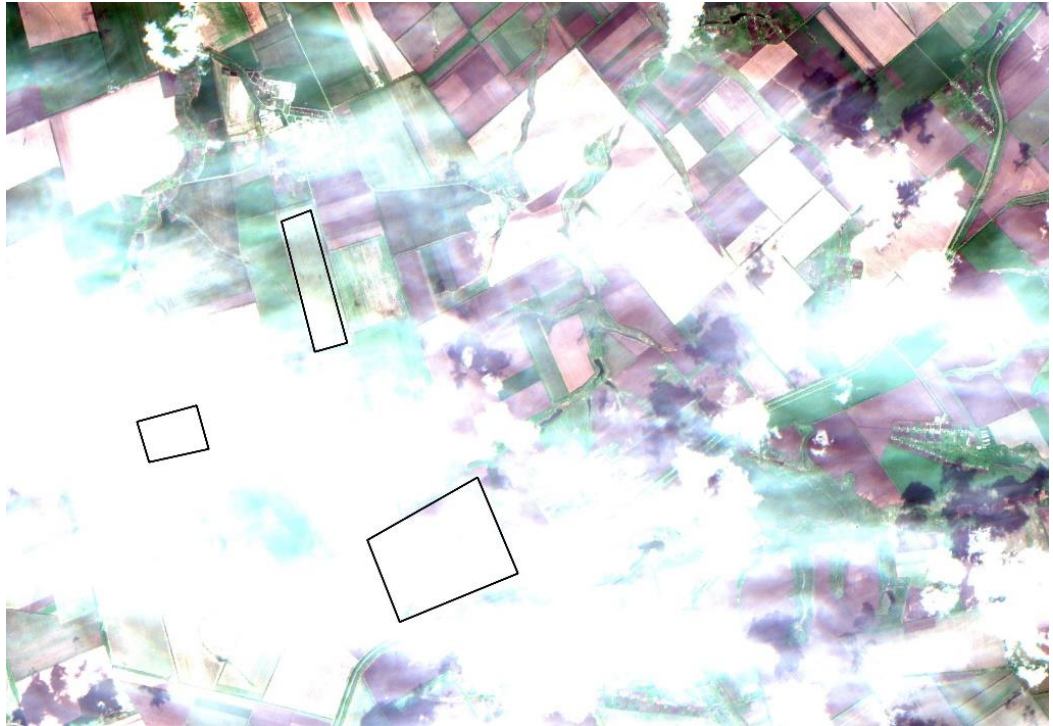


Рис. 1.9 Знімок Sentinel-2 для Київської області

- невизначеність (неточність) в даних у вибірці, обумовлена її підготовкою на основі фотоінтерпретації (суб'єктивність думки експерта (рис. 1.9));

- перекривання декількох зображень для однієї ділянки поверхні Землі;
- різне часове та просторове розрізнення – отриманих даних одного супутника може бути не достатньо (наприклад, оптичні знімки можуть бути захмарені (рис. 1.10).



Sentinel-2 (22.04.2016)



Landsat-8 (13.02.2016)

Рис. 1.10 Супутникові продукти із різним просторовим та часовим розрізненням

Для подолання даних проблем необхідно об'єднувати дані різних супутників – виконувати злиття даних. Різні методи будуть запропоновані для розв'язання цієї комплексної задачі в наступних розділах.

РОЗДІЛ 2

МЕТОДИ КАРТОГРАФУВАННЯ ЗЕМНОГО ПОКРИВУ НА ОСНОВІ ЗЛИТТЯ ДАНИХ

2.1 Загальна схема класифікації земного покриву для великих територій на основі даних високого розрізнення

Для розв'язання задачі побудови карти земного покриву високого розрізнення для великих територій пропонується підхід, що базується на використанні трьох різних парадигм з різних областей комп'ютерних наук: принципу «розділяй і володарюй» з теорії алгоритмів [158], методу активного навчання з теорії машинного навчання [159, 160] і методу відновлення зображень з області комп'ютерної обробки цифрових зображень [161]. Використовуючи різномірну геопросторову інформацію для побудови карти земного покриву (супутникові дані високого розрізнення, як оптичні так і радарні, зібрані дані наземних досліджень (або отримані шляхом фотоінтерпретація по супутникових даних)) зустрічаємося із низкою проблем (детально наведено в першому розділі, пункт 1.7). Для їх вирішення пропонується використовувати підхід (обґрунтований у першому розділі, пункт 1.5), що базується на основі методів злиття даних та складається із наступних етапів (рис. 2.1):

- збір та попередня обробка даних;
- класифікація, побудова результуючої карти, покращення отриманих результатів класифікації (постобробка);
- візуалізація.

Для побудови карти земного покриву для великої території будемо використовувати безкоштовні супутникові дані (як радарні, так і оптичні) із високим просторовим розрізненням (Landsat (30 м), Sentinel (10 м) та інші).

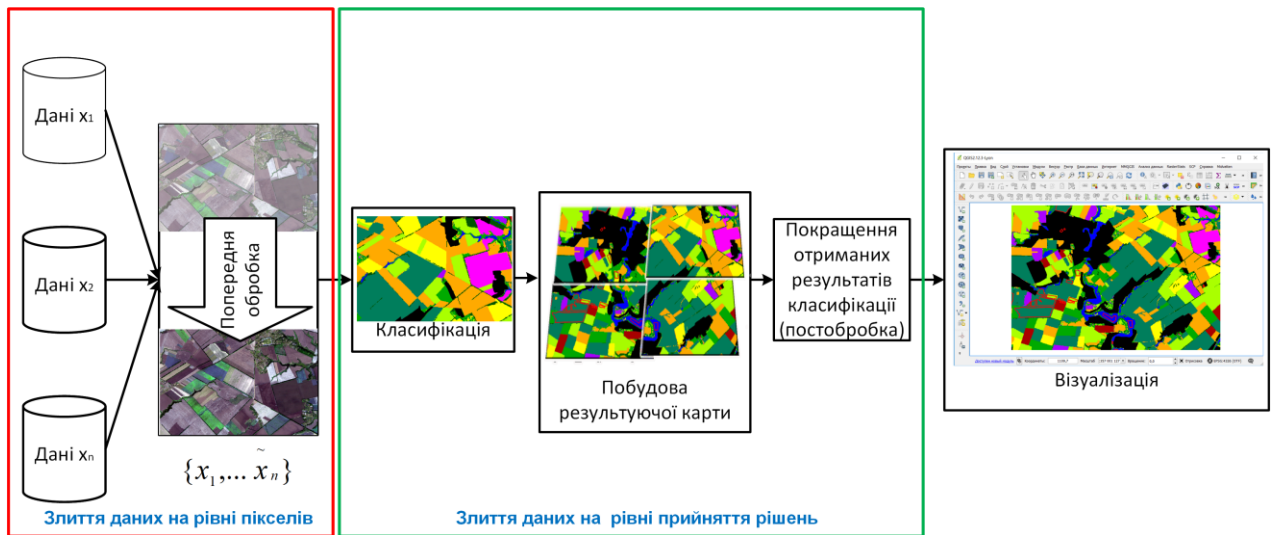


Рис. 2.1 Загальна схема методу картографування земного покриття

На основі одного знімка неможливо отримати достовірну карту, оскільки оптичні та радарні дані мають свої недоліки (які були раніше описані в першому розділі, пункт 1.2). Тому в якості вхідних даних для класифікації пропонується використовувати часовий ряд даних, отриманих з усіх доступних супутників. Оскільки ці дані мають різне просторове розрізнення і різну природу, то вони потребують різної попередньої обробки. Таким чином для забезпечення попередньої обробки пропонується виконувати злиття даних на рівні пікселів. Отже, на першому етапі попередньої обробки даних здійснюється радіометрична та геометрична корекція, ортотрансформування, атмосферна корекція (лише для оптичних даних), фільтрація знімків (лише для радарних даних), перепроєктування даних до однієї проекції, зведення їх до одного просторового розрізнення та нормалізація. При цьому пропонується розширити вектор інформативних ознак для подальшої класифікації шляхом злиття різночасових даних різної природи на рівні пікселів (оптичних та радарних, причому оптичні мають значення пікселів, яких не було в місцях, де дані були захмарені (за рахунок відновлення пропусків даних)).

Наступним етапом картографування земного покриття є класифікація даних (використовується номенклатура класів LUCAS, розглянута в першому розділі). В якості метода класифікації пропонується використовувати ансамбль

нейронних мереж на основі багатошарових персептронів (MLP – multi-layer perceptron), в якому кількість прихованих нейронів підбирається в залежності від кількості каналів зображень, поданих на навчання (обґрунтування вибору метода наведено в пункті 1.4 першого розділу).

Кожна нейронна мережа після класифікації видає апостеріорну ймовірність належності вхідного пікселя до певного класу. В ансамблі рахується середня апостеріорна ймовірність виходів усіх мереж. Використовуючи отримані ймовірності пропонується проводити злиття даних на рівні прийняття рішень для визначення належності кожного пікселя до певного класу, результатом якого є карта класифікації для кожної із сцен.

Як було обґрунтовано в першому розділі (пункт 1.2), неможливо отримати точну карту класифікації для великої території, використовуючи один знімок. Тому, проводиться класифікація по всіх можливих знімках, щоб покрити досліджувану територію. Результатом попиксельної класифікації є набір даних, де кожному пікселю відповідає певний клас земного покриття. Враховуючи такий підхід, виникає необхідність злиття результатів класифікації для побудови цілісної карти (рис. 2.8). Для цього пропонується розробити методи злиття даних на рівні прийняття рішень, що базуються на використанні інформації про достовірність кожного пікселя на початковому та кінцевому етапах обробки (захмареність, апостеріорна ймовірність, точність класифікації). В результаті отримаємо карту земного покриття для досліджуваної території. При цьому, часто можуть виникати помилки класифікації, коли в межах одного полігона будуть пікселі, що належать різним класам, так як класифікація виконується попиксельно, що потребує додаткового етапу постобробки (рис. 3.14).

Для покращення карти земного покриття пропонується використовувати методи злиття растрової (карта класифікації) та векторної (містить межі полігонів) інформації [162]. У результаті в межах кожного полігона будуть пікселі, що відносяться до одного класу.

2.2 Злиття даних на рівні пікселів на етапі попередньої обробки

На сьогоднішній день існує багато космічних апаратів, які надають супутникові дані в різному вигляді (просторове розрізнення, рівень попередньої обробки супутникових даних, географічна прив'язка і тд.). Для того, щоб на основі таких різнорідних даних побудувати карту земного покриття для великої території необхідно провести їх попередню обробку та злиття до гомогенного виду.

Злиття даних різних супутників за різні часові дати відбувається в кілька етапів і являє собою: отримання нових даних, їх попередню обробку, злиття і збереження в уніфікованому форматі. Даний функціонал наведений у вигляді схеми на рис. 2.2. Для кожного типу даних реалізований окремий процес обробки.

В межах даного процесу здійснюється отримання знімків з оптичних $X_L = \{X_i\}_{i=1}^k$, та радіолокаційних $X_R = \{X_i\}_{i=1}^k$ приладів, що необхідні для використання на досліджуваній території. Можливі кілька варіантів завантаження даних (завантаження даних за конкретну дату чи за діапазон дат). Після отримання завантажених даних (як оптичних так і радіолокаційних), що покривають досліджувану територію, виконується їх попередня обробка. Процес виконання завершується у випадку, коли закінчуються завантажені знімки.

При попередній обробці оптичних даних виконуються наступні етапи попередньої обробки:

- для виявлення хмар та тіней від них пропонується використовувати алгоритм Fmask [22]:

$$\overline{X}_L = \{f_{fmask}(X_L, T) \cup X_L\}; \quad (2.1)$$

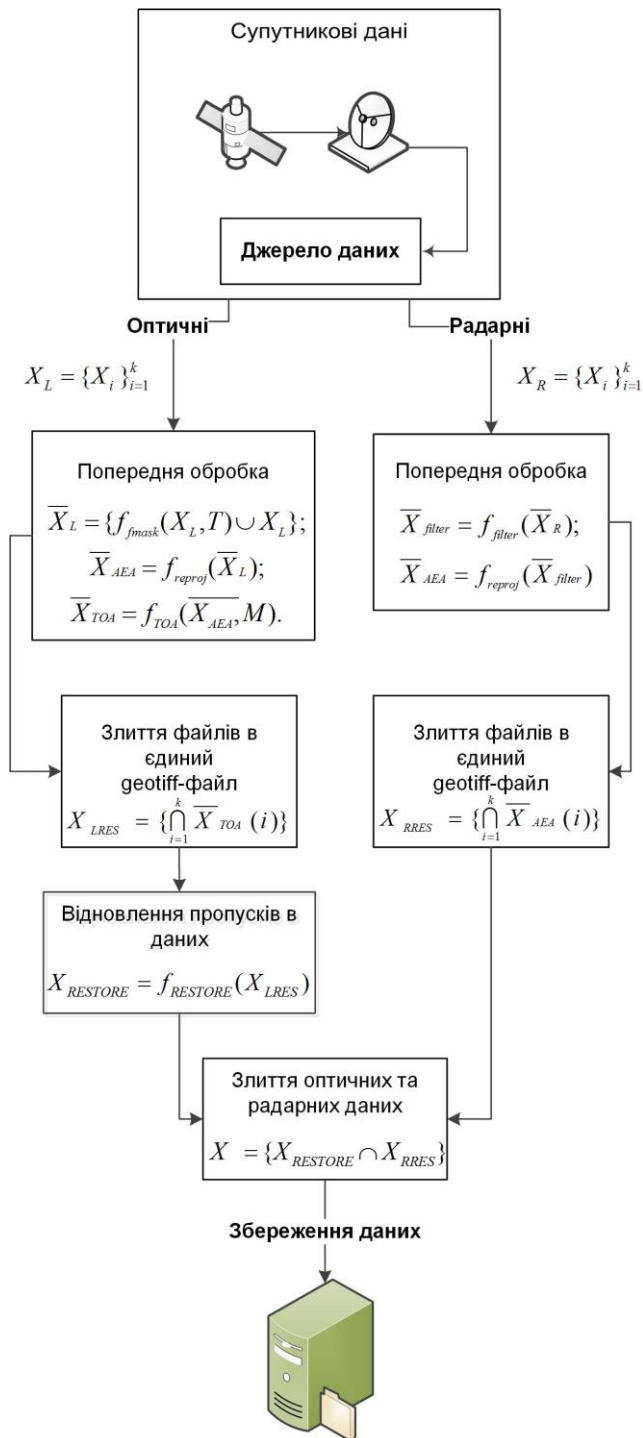


Рис. 2.2 Злиття даних різної та однорідної природи за різні часові дати

- кожне супутникове зображення перепроєктовується у проекцію Albers Equal Area (AEA):

$$\overline{X}_{AEA} = f_{reproj}(\overline{X}_L); \quad (2.2)$$

- радіометрична та геометрична корекція, ортотрансформування, атмосферна корекція. Дані кроки виконуються із використанням продуктів для оптичних знімків (залежно від супутника):

$$\overline{X}_{TOA} = f_{TOA}(\overline{X}_{AEA}, M). \quad (2.3)$$

Для попередньої обробки радарних даних виконуються наступні кроки:

- перетворення цифрових значень (DN) до коефіцієнта зворотнього розсіювання (σ_0) з використанням коефіцієнтів з файлу метаданих продукту;
- фільтрація знімків за допомогою віконного фільтра:

$$\overline{X}_{filter} = f_{filter}(\overline{X}_R); \quad (2.4)$$

- ортотрансформування за допомогою SRTM DEM;
- кожне супутникове зображення перепроєктовується у проекцію Albers Equal Area (AEA):

$$\overline{X}_{AEA} = f_{reproj}(\overline{X}_{filter}); \quad (2.5)$$

- зміна просторового розрізнення зображень.

Всі ці кроки попередньої обробки радарних даних виконуються за допомогою The Next ESA SAR Toolbox (NEST) програмного забезпечення з відкритим вихідним кодом [163].

Наступним етапом обробки супутникових даних є процес злиття всіх каналів як оптичних $X_{LRES} = \{\bigcap_{i=1}^k \overline{X}_{TOA}(i)\}$, так і радіолокаційних $X_{RRES} = \{\bigcap_{i=1}^k \overline{X}_{AEA}(i)\}$, у два загальні файли. Основним аспектом є те, що всі вхідні зображення повинні бути в однаковій системі координат. Після обробки даних вони зберігаються як растровий знімок, що містить всі канали, які покривають досліджувану територію. Таким чином відбувається злиття часового ряду супутникових даних (окремо радіолокаційних та оптичних) в один вектор (рис. 2.3).

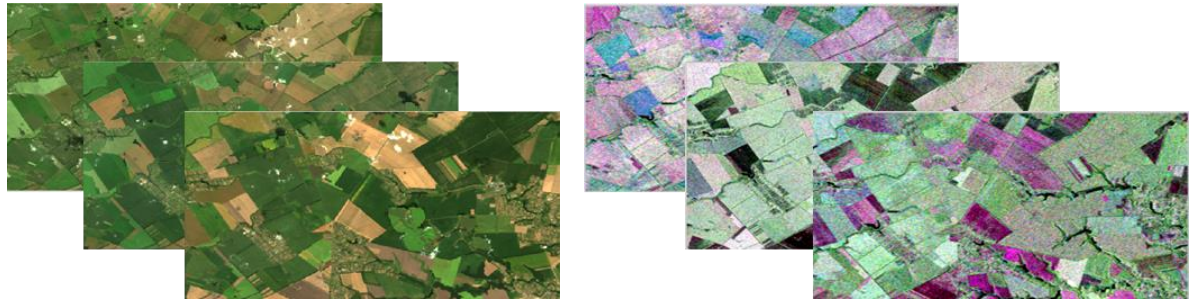


Рис. 2.3 Часовий ряд оптичних та радіолокаційних даних

Наступним етапом обробки оптичних знімків є відновлення пропусків в даних у випадку їхньої значної захмареності, використовуючи часовий ряд даних. Ідея цього підходу полягає в послідовному застосуванні нейронних мереж різного типу для відновлення невизначеностей (хмар на знімках). Процедура відновлення виконується згідно розробленого методу на основі самоорганізованих карт Кохонена (SOM), детально описаної в [164 – 168]:

$$X_{RESTORE} = f_{RESTORE}(X_{LRES}). \quad (2.6)$$

Отримані оптичні та радарні знімки об'єднуються в один файл:

$$X = \{X_{RESTORE} \cap X_{RRES}\}. \quad (2.7)$$

Запропонований підхід полягає у розробці методів обробки часових рядів оптичних та радіолокаційних супутникових даних з метою забезпечити злиття даних на рівні пікселів за рахунок формування узагальненого вектора ознак з урахуванням «пропусків» у оптичних даних.

Таким чином, після обробки зберігається результуючий файл в форматі GeoTIFF, що містить набір значень всіх радарних та оптичних знімків, причому оптичні мають значення пікселів, яких не було в місцях, де дані були захмарені (за рахунок відновлення пропусків даних (2.6)).

Отриманий GeoTIFF використовується для побудови карти земного покриття.

2.3 Побудова карт земного покриття для великих територій на основі методів злиття даних

Для побудови карти земного покриття великої території на основі даних з високим просторовим розрізненням використовується два підходи. Перший полягає в класифікації кожного знімка окремо з подальшим злиттям результатів класифікації в загальну карту для даної території (злиття на рівні прийняття рішень). Другий підхід полягає в класифікації на основі даних всіх супутникових зображень (часового ряду зображень), що покривають протягом року (вегетаційного періоду) дану територію [53, 163].

Використання знімків високого розрізнення (особливо їх часового ряду) займає достатньо великий обсяг пам'яті. Тому, для мінімізації вимог до обчислювальних ресурсів системи при побудові карти земного покриття велику територію (наприклад, Україну) пропонується розбити на області, що не перетинаються (згідно з принципом «розділяй і володарюй») та розв'язувати задачу класифікації для кожної з них окремо. У загальному випадку це можна здійснити двома способами:

- розбиття на області однакового розміру з урахуванням покриття супутниковими даними [169];
- розбиття з урахуванням адміністративних областей.

Другий спосіб розбиття не забезпечує рівність площ елементарних областей, однак він має наступну перевагу: дозволяє залучити незалежну інформацію (наприклад, офіційні статистичні дані, які доступні для адміністративних областей) для валідації отриманих карт. Слід зазначити, що саме такий спосіб був використаний при створенні карти класів поверхні Землі CORINE, коли кожна країна забезпечувала дослідження своєї власної території [170].

Проте, окрім супутникових даних необхідною складовою є навчальна та тестова вибірки, формування яких детально розглянемо у наступному пункті.

2.3.1 Метод формування тестової та навчальної вибірки для побудови карти земного покриття

Запропонуємо загальний підхід до формування навчальної та тестової вибірок для побудови карти земного покриття великої території. Одним із шляхів формування достовірної вибірки є проведення наземних досліджень. У випадку, якщо дослідження неможливі, то вибірка формується шляхом фотоінтерпретації (наприклад, для побудови вибірки за ретроспективними даними). Будемо вважати, що при наземних дослідженнях або в процесі фотоінтерпретації зібрані дані представлені у вигляді геометричних полігонів.

Нехай P — множина меж полігонів, G — вибірка для деякої території, що складається з пар «межа полігону» і відповідний даному полігону клас земної поверхні

$$G = \{(p, k) \mid p \in P, k = \overline{1, K}\}, \quad (2.8)$$

де k — клас земної поверхні, K — загальна кількість класів. Вибірку доцільно зберігати у векторному вигляді для можливості її використання при класифікації супутникових зображень різного просторового розрізнення і розміру. Кількість полігонів у вибірці (потужність множини) позначимо через $|G|$. Вибірку розділимо на навчальну і тестову приблизно в рівних пропорціях, як це прийнято при використанні індуктивного підходу. При цьому:

$$G = G_{train} \cup G_{test}, \quad (2.9)$$

$$G_{train} \cap G_{test} = \emptyset. \quad (2.10)$$

В якості вхідних даних в задачі побудови карт земного покриття будемо розглядати часовий ряд супутникових зображень, на яких пікселі, вкриті хмарами, відновлені методом [161]. На основі часового ряду супутникових зображень для кожного пікселя обраної території формується вектор:

$$\mathbf{x} = (x_1, x_2, \dots, x_n) \in R^n, \quad (2.11)$$

$$n = i \cdot c, \quad (2.12)$$

де i — кількість супутникових зображень, що покривають даний піксель протягом року, s — кількість різних спектральних каналів супутникового зображення (детальніше щодо формування вхідного вектора супутникових даних на основі методу злиття даних на рівні прийняття рішень описано в пункті 2.2 другого розділу дисертаційного дослідження).

Таким чином, на основі інформації зі спектральних каналів супутникового зображення можна сформувати множину пікселів вибірки, де кожному пікселю, що належить множині полігонів G , відповідає номер класу

$$T(G) = \{(\mathbf{x}_i, k_i), i = \overline{1, N}\}, \quad (2.13)$$

де N — кількість всіх пікселів у вибірці. Відзначимо, що потужність і склад даної множини залежить від набору полігонів G . Відповідно із розбиттям множини полігонів G (формули 2.9-2.10), множина пікселів T теж розбивається випадковим чином на дві незалежні підмножини для навчання класифікатора і його тестування:

$$T = T_{train} \cup T_{test}, \quad (2.14)$$

$$T_{train} \cap T_{test} = \emptyset. \quad (2.15)$$

Таким чином, отримано тестову та навчальну вибірки, що будуть використовуватись для побудови карти земного покриття.

2.3.2 Ансамблевий метод класифікації

Після попередньої обробки даних отриманий єдиний вектор супутникових даних та сформовані тестова та навчальна вибірки, які подаються на вхід класифікатора. Для класифікації пропонується застосовувати багатошарові нейронні мережі персептронного типу (обґрунтування використання нейромережевого підходу наведено у першому розділі, пункт 1.4). Для підвищення ефективності класифікації доцільно застосовувати не один окремих класифікатор, а їх ансамбль (практично доведено у третьому розділі, пункт 3.1). При цьому під час класифікації відбувається злиття даних на рівні прийняття рішень (рис. 2.4).

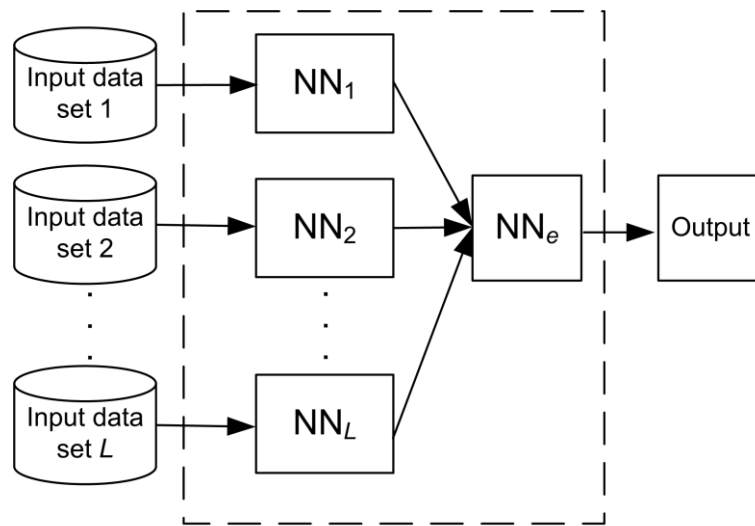


Рис. 2.4 Реалізація класифікації: злиття даних відбувається на рівні прийняття рішень

Під час навчання нейронної мережі з використанням багатошарового персептрона функція помилок крос-ентропії зводиться до мінімуму:

$$E(\mathbf{w}) = -\ln p(\mathbf{T} | \mathbf{w}) = -\sum_{n=1}^N \sum_{k=1}^K t_{nk} \ln y_{nk} \rightarrow \min, \quad (2.16)$$

де $\mathbf{w}$ — вектор вагових коефіцієнтів, $\mathbf{T}$ — множина, яка складається з вектору спектральних каналів для кожного пікселя (2.11) та відповідного йому класу k в навчальній вибірці; N — кількість полігонів у вибірці; K — число класів; t_{nk} — цільові виходи; y_{nk} — виходи MLP, $p(\cdot)$ — умовна ймовірність.

На виході кожного класифікатора отримуємо апостеріорну ймовірність належності вхідного образу до певного класу. Виходи окремих класифікаторів об'єднуються простим додаванням, після чого обчислюється середня ймовірність належності вхідного образу конкретному класу (рис. 2.5).

Вхідний образ відносять до класу з максимальною апостеріорною ймовірністю:

$$k^* = \arg \max_{k=1, K} p_k^e, \quad p_i^e = \frac{1}{L} \sum_{l=1}^L p_i^l, \quad (2.17)$$

де k^* — клас, до якого відносить вхідний образ ансамбль класифікаторів, p_i^e —

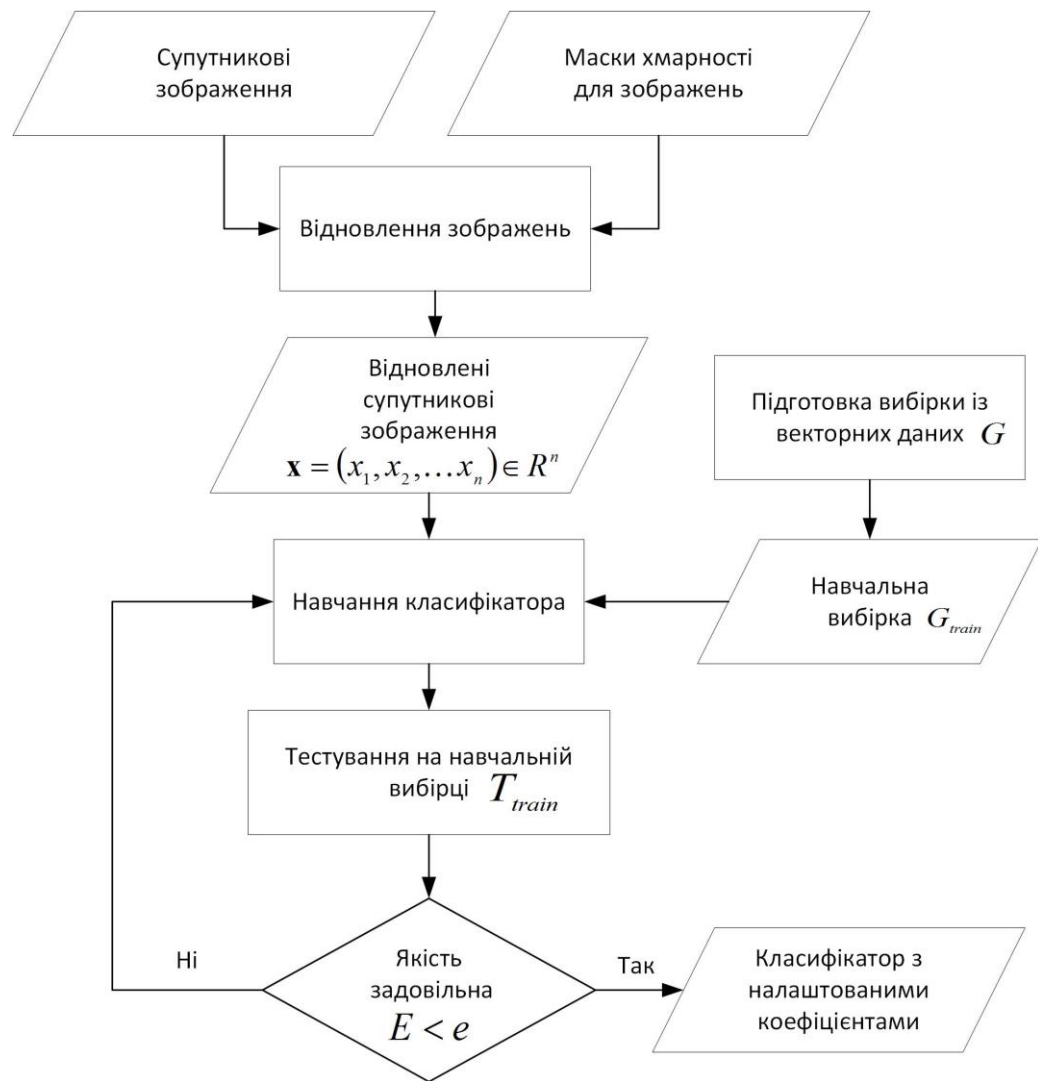


Рис. 2.6 Навчання класифікатора або налаштування вагових коефіцієнтів MLP

Після налаштування параметрів класифікатора виконується попіксельна класифікація зображення. В результаті виконання такої операції для всіх пікселів із досліджуваної області, отримуємо карту класифікації. Точність отриманої карти класифікації будемо оцінювати на незалежній тестовій вибірці (формування вибірок описано в пункті 2.3.1), а також шляхом порівняння з даними офіційної статистики.

Для оцінки точності класифікації на незалежній вибірці будемо використовувати наступні позначення та формули. Кількість пікселів n_{ij} у тестовій вибірці, що реально належать класу i , але помилково віднесені до класу j , позначимо n_{i+} :

$$n_{i+} = \sum_{j=1}^K n_{ij}, \quad (2.18)$$

Число пікселів n_{ij} , що реально належать класу j , але помилково віднесені до класу i позначимо n_{+j}

$$n_{+j} = \sum_{i=1}^K n_{ij}. \quad (2.19)$$

Тоді загальна точність карти класифікації визначається:

$$OA = \frac{\sum_{k=1}^K n_{kk}}{N}. \quad (2.20)$$

З урахуванням введених позначень можна обчислити помилки першого і другого роду, або точність «виробника» PA і точність «користувача» UA , а також індекс Каппа K , що ілюструє рівень «випадковості» класифікації [165]:

$$PA = \frac{n_{jj}}{n_{+j}}, \quad (2.21)$$

$$UA = \frac{n_{ii}}{n_{i+}}, \quad (2.22)$$

$$K = \frac{N \cdot \sum_{i=1}^K n_{ii} - \sum_{i=1}^K n_{i+} \cdot n_{+i}}{N^2 - \sum_{i=1}^K n_{i+} \cdot n_{+i}}. \quad (2.23)$$

Нульове значення індексу Каппа показує, що результат класифікації так само узгоджується з вибіркою, як і випадковим чином призначені класи.

Якщо якість (точність класифікації на незалежній тестовій вибірці) отриманої карти задовільна, вона приймається як результуюча карта класифікації. Якщо ж точність не задовільна, виконується донавчання або вибір нового класифікатора (активне навчання).

На рис. 2.7 наведена узагальнена блок-схема роботи класифікатора з донавчання (активне навчання), у разі незадовільної точності карти класифікації на незалежній тестовій вибірці.

Описаний підхід, що включає відновлення зашумлених пікселів і

активне навчання, забезпечує автоматичну побудову карти класифікації для кожної обраної частини території (наприклад, адміністративної області України).

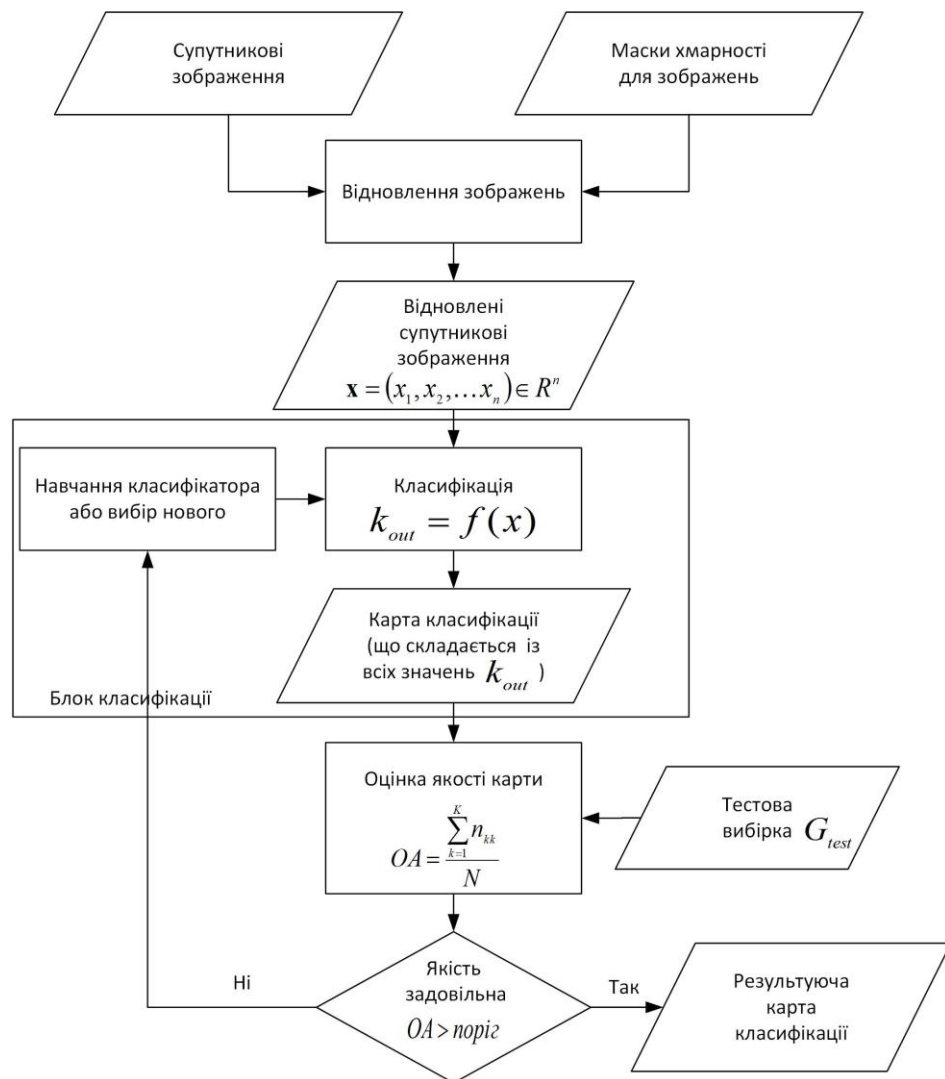


Рис 2.7 Алгоритм активної класифікації супутникових зображень

Одним із головних факторів, що впливають на точність побудованої карти земного покриття є якість вибірки. Можливості покращення тестової та навчальної вибірок наведені у наступному пункті.

2.3.4 Покращення тестової та навчальної вибірок в процесі побудови карти земного покриття

Постановку задачі покращення вибірки при побудові карти земного

покриву можна сформулювати наступним чином: необхідно ітеративно модифікувати множину G (описана в пункті 2.3.1), щоб виконувалися наступні умови

$$\begin{aligned} E(T_{train}) &\longrightarrow \min, \\ UA(G_{test}) &\longrightarrow \max, \\ PA(G_{test}) &\longrightarrow \max, \end{aligned} \quad (2.24)$$

$$|G| > g_{critical}, \quad (2.25)$$

де E – функція помилки нейронної мережі (2.16), при умовах (2.9-2.10), $g_{critical}$ — мінімальна кількість полігонів, необхідних для задовільного навчання нейронної мережі [171, 172].

Якщо навчальна і тестова вибірки формуються на основі фотоінтерпретації, то не можна стверджувати, що вони містять абсолютно достовірну інформацію (вплив суб'єктивного фактора). Тому класифікацію необхідно виконувати ітераційно, при необхідності редагуючи і фільтруючи вибірки на кожному етапі. Вибірка ітеративно модифікується шляхом усунення помилок (модифікації меж полігонів) і додаванням нових полігонів до тих пір, поки якість класифікації не стане задовільною. Причому кількість обраних полігонів відповідає вимогам репрезентативності представлених класів та рівномірного покриття полігонами досліджуваної території.

Після побудови карти земного покриття необхідно зіставити її з тестовою вибіркою G_{test} і побудувати матрицю невідповідності за результатами класифікації. Оскільки розбіжності можуть виникати за різних причин, то в першу чергу необхідно забезпечити виконання умови (2.25), тобто репрезентативність тестової вибірки для всіх класів, а також рівномірний розподіл навчальної вибірки G_{train} по досліджуваній території. Для якісної класифікації також необхідно, щоб класи в навчальній вибірці були розподілені пропорційно відповідним площам на досліджуваній території. Не враховуючи цих вимог буде отримана висока точність класифікації на тестовій вибірці $UA(G_{test})$, $PA(G_{test})$, однак результуюча карта класифікації виявиться

незадовільною. Експерт повинен проаналізувати результати класифікації даного поля, вибірку G і часовий ряд супутникових зображень (2.11). Найбільш поширеними причинами невідповідності карти класифікації тестовій вибірці є наступні артефакти вибірки: перекриття полігонів, що відносяться до різних класів, сумнівні полігони (експерту складно ідентифікувати клас певного полігону), «змішані» полігони (що містять вкраплення інших класів, найчастіше невеликих озер або невеликі групи дерев), нетипові полігони (полігон, віднесений до якого-небудь класу, суттєво відрізняється від інших представників цього ж класу) та інші проблеми, пов'язані з якістю супутникових зображень. Для забезпечення коректного навчання нейронної мережі сумнівні полігони видаляються з вибірки, неправильно класифіковані полігони додаються до правильного класу, нетипові полігони вважаються винятками з правил, тобто викидаються, і видаляються. Ітераційний процес завершується після ідентифікації всіх невідповідностей у вибірці. При цьому отриману карту класифікації можна розглядати як остаточний результат.

2.4 Злиття даних на рівні прийняття рішень для отримання результуючої карти земного покриття

Для побудови карт земного покриття для великих територій (на рівні областей або країн) з використанням супутникових даних, виникає проблема, коли неможливо покрити досліджувану територію одним зображенням середнього або високого просторового розрізнення (рис. 2.8). Причиною цього може бути велика площа досліджуваної території, відсутність даних з одного космічного апарата, наявність на супутникових зображеннях хмарності, похибок, тощо.

Для побудови результуючої карти пропонується три методи злиття даних на рівні прийняття рішень.

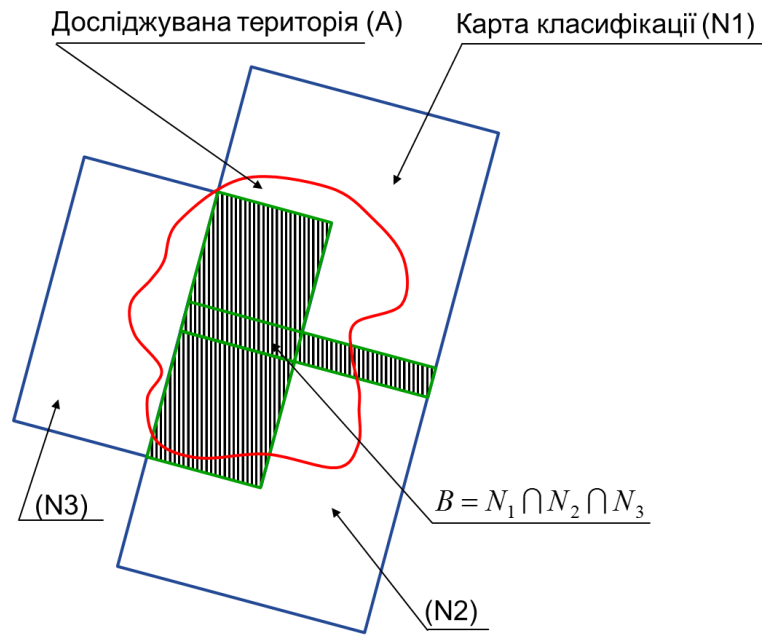


Рис. 2.8. Неузгодженість результатів класифікації

Перший метод злиття полягає в попіксельному прийнятті рішення на основі ймовірності приналежності даного пікселя певному класу на кожному зображенні, що покриває даний піксель:

$$\text{Im}_{\text{res}}(i, j) = \begin{cases} \text{Im}_1(i, j), P(\text{Im}_1(i, j)) > P(\text{Im}_2(i, j)) \\ \text{Im}_2(i, j), P(\text{Im}_1(i, j)) \leq P(\text{Im}_2(i, j)) \end{cases}, \quad (2.26)$$

де $\text{Im}_1(i, j)$, $\text{Im}_2(i, j)$ - результати класифікації пікселя (i, j) по кожному окремому знімку, $\text{Im}_{\text{res}}(i, j)$ - остаточна класифікація, $P(\cdot) = UA(\cdot)$, згідно (2.22), де $P(\cdot)$ — ймовірність події із матриці невідповідності (точність користувача).

Другий метод злиття на рівні прийняття рішень зводиться до наступного: рішення про приналежність даного пікселя деякому класу приймається з урахуванням якості знімків, що його покривають, тобто маски відновлених пікселів, оскільки результат класифікації відновлених пікселів не відрізняється високою надійністю:

$$\text{Im}_{\text{res}}(i, j) = \begin{cases} \text{Im}_1(i, j), \text{Mask}(\text{Im}_1) < \text{Mask}(\text{Im}_2) \\ \text{Im}_2(i, j), \text{Mask}(\text{Im}_1) \geq \text{Mask}(\text{Im}_2) \end{cases}, \quad (2.27)$$

де $\text{Im}_1(i, j)$, $\text{Im}_2(i, j)$ — результати класифікації пікселя (i, j) по кожному окремому знімку, $\text{Im}_{\text{res}}(i, j)$ — остаточний результат класифікації, $\text{Mask}(\text{Im}_i)$ —

кількість відновлених пікселів на зображенні, результатом класифікації якого є Im_i , $i = \overline{1,2}$.

У межах *третього методу* пропонується використовувати класифікацію на основі часових рядів супутникових даних, а для отримання результуючої карти застосовується метод злиття даних на рівні прийняття рішень із урахуванням апостеріорної ймовірності належності кожного пікселя певному класу та маски хмарності:

$$\text{Im}_{\text{res}}(i, j, k) = \begin{cases} \text{Im}_{11}(i, j, k), P(\text{Im}_{11}(i, j, k)) > P(\text{Im}_{22}(i, j, k)) \\ \text{Im}_{22}(i, j, k), P(\text{Im}_{11}(i, j, k)) \leq P(\text{Im}_{22}(i, j, k)) \end{cases} \quad (2.28)$$

де $\text{Im}_{11}(i, j, k)$, $\text{Im}_{22}(i, j, k)$ — результати класифікації пікселя (i, j) на основі часового ряду, $\text{Im}_{\text{res}}(i, j, k)$ — остаточно класифікація, k — кількість захмарених пікселів на зображенні, $P(\cdot) = p_i^e$, згідно (2.17), де $P(\cdot)$ — апостеріорна ймовірність, отримана при класифікації.

2.4.1 Метод злиття даних із різним просторовим розрізненням

Розглянемо процес аналізу стану деякої території, що покривається даними декількох супутників. Прикладом такої задачі може бути оцінка пошкоджень посівів внаслідок посухи (детальне застосування розглянуто в п'ятому розділі).

Розглянемо процес злиття даних декількох супутників $d, d = \overline{1, D}$, де D - кількість супутників.

Нехай маємо деяке поле A , яке складається з пікселів, які покриваються кількома знімками супутників з різним просторовим розрізненням (рис. 2.9). Щоб визначити, до якого класу належить поле, підрахуємо кількість пікселів, що відноситься до кожного класу з урахуванням даних різних супутників. Тоді, для поля A присвоюємо той клас, в якому найбільше пікселів в межах поля A .

Розглянемо більш детально процес класифікації з використанням різнорідних даних.

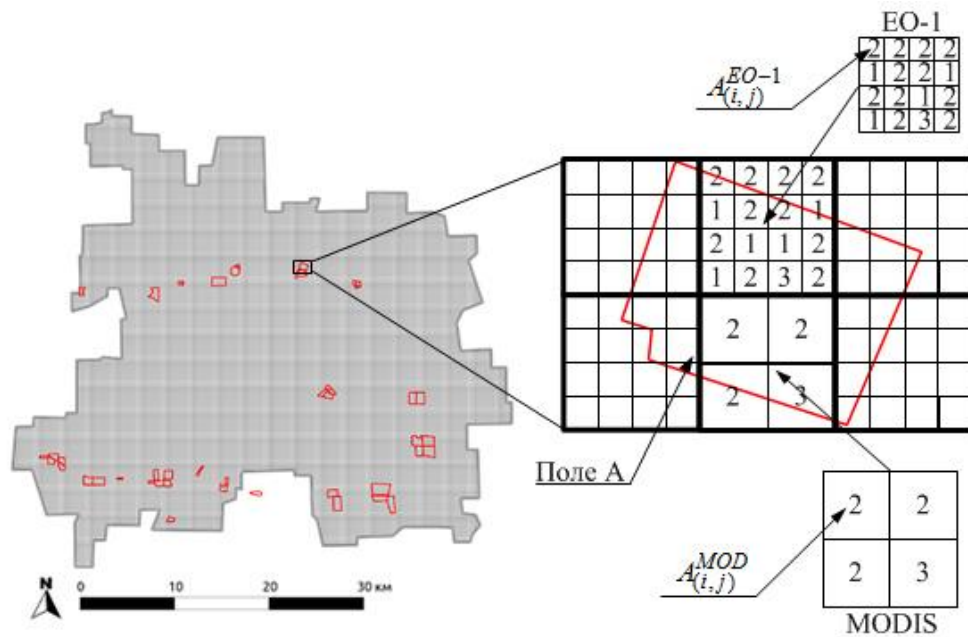


Рис. 2.9 Злиття даних із використанням даних із різних джерел

Нехай на карті класифікації L класів. Для поля A на карті класифікації, побудованої по супутнику $d, d = \overline{1, D}$, кожний піксель характеризується дискретними координатами (i, j) та класом, якому належить піксель — $C_{(i,j)}^{(d)}$.

$$A_{(i,j)}^{(d)} = \{(i, j), C_{(i,j)}^{(d)}\}. \quad (2.29)$$

Кількість пікселів поля A , віднесених на карті до кожного з класів $l, l = \overline{1, L}$, можна визначити наступним чином:

$$K_l^{(d)} = \sum_{(i,j) \in A} 1 (C_{(i,j)}^{(d)} = l). \quad (2.30)$$

Виконаємо злиття карт класифікації, побудованих для різних супутників. Загальне число пікселів поля A , віднесених на кожній карті $d, d = \overline{1, D}$ до кожного з класів $l, l = \overline{1, L}$, можна визначити наступним чином:

$$K_l = \sum_{d=1}^D W_d \cdot K_l^{(d)}, \quad l, l = \overline{1, L}, \quad (2.31)$$

де W_d — ваговий коефіцієнт, що враховує просторове розрізнення різних супутників. Для визначеності, з урахуванням специфіки зображень, пропонується визначати W_d із формули (2.31) наступним співвідношенням:

$$W_d = \left[\frac{S_d}{\min_{\alpha} S_d} \right], \quad d = \overline{1, D}, \quad (2.32)$$

де S_d – площа пікселя кожного з супутників $d = \overline{1, D}$, α – кількість всіх супутників, які використовуються для визначення вагових коефіцієнтів, $[\cdot]$ – операція обчислення цілої частини.

Поле A віднесемо до того класу c , до якого відноситься максимальна кількість пікселів, що належать цьому полю, згідно зваженої суми (2.31).

$$c = \arg \max_i K_i. \quad (2.33)$$

2.4.2 Методи покращення карти земного покриття на основі злиття геопросторових даних

Побудова карт земного покриття на основі супутникових даних, як правило, зводиться до розв'язання задачі попінкельної класифікації знімків та постобробки результатів для зменшення фрагментації карти. Дуже часто виникає ситуація, коли в межах одного поля ідентифіковані пікселі, що відносяться до різних класів. Для підвищення правдоподібності результатів доводиться приймати рішення, до якого із класів віднести всі пікселі поля (полігона). Причому кількість класів, ідентифікованих в межах одного полігона, може бути більше двох, а кількість пікселів кожного класу близька між собою. Для вирішення таких проблем застосовують методи злиття даних на рівні прийняття рішень (data fusion) [5].

Основна ідея дослідження полягає в тому, що проводиться аналіз кожного полігона згідно меж для визначення приналежності його до конкретного класу. Пропонується два методи злиття даних [48]. Перший метод злиття растрової та векторної інформації базується на методі голосування – полігону присвоюється той клас, пікселів якого виявилось найбільше:

$$K_{out}^i = \arg \max_j \left(\sum_{j=1}^K k_{out}^j \right), \quad (2.34)$$

де K_{out}^i – значення класу для i -го полігону,

$$k_{out}^j(x, y) = \begin{cases} 0, & k_{out}(x, y) \neq j \\ 1, & k_{out}(x, y) = j \end{cases}, (x, y) \in p^i. \quad (2.35)$$

Зауважуючи, що для отримання цієї карти класифікації використовуємо метод відновлення відсутніх значень на знімках, пропонується враховувати ці відновлені дані, як менш надійні з певними штрафними ваговими коефіцієнтами (зважений метод голосування). Другий метод враховує кількість захмарених знімків, що використовувались під час класифікації. Таким чином, більш надійними вважаються результати класифікації для найменш захмарених пікселів.

Для призначення штрафних ваг враховуємо маску хмарності (наприклад, для території Київської області), на якій показана кількість незахмарених знімків для кожного пікселя $m(i, j)$ (рис. 2.10).

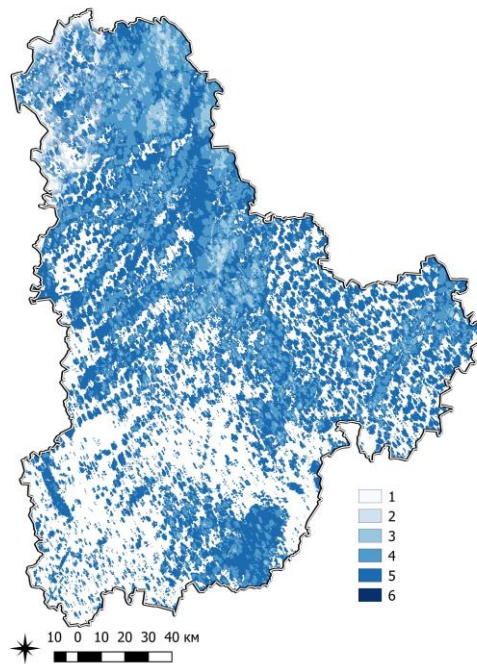


Рис. 2.10 Маска хмарності для Київської області

Тоді штрафні ваги обчислюємо за формулою $w(i, j) = \begin{cases} \frac{1}{m(i, j)^2}, & m(i, j) \neq 0 \\ 0, & m(i, j) = 0 \end{cases}$, і

метод голосування реалізується на основі співвідношення:

$$K_{out}^i = \arg \max_j \left(\sum_{j=1}^K k_{out}^j \cdot w(x, y) \right), (x, y) \in p^i. \quad (2.36)$$

Результати проведеного дослідження показали, що метод із врахуванням захмареності даних дає вищу загальну точність карти в порівнянні із методом голосування. Детальніше результати методів покращення карт класифікації представлені у наступному розділі.

2.5 Висновки до розділу

У другому розділі задача класифікації земного покриття для великих територій на основі даних високого розрізнення розглянута як задача обробки великих обсягів геопросторових даних, які можуть містити невизначеності (наприклад, області, покриті хмарним покритвом). Для її вирішення запропоновано використовувати три різних парадигми комп'ютерних наук, а саме, метод декомпозиції («розділяй і володарюй») з теорії алгоритмів, активного навчання із сфери інтелектуальних обчислень і методу відновлення супутникових знімків з області комп'ютерної обробки цифрових зображень. Комплексне застосування цих трьох складових дозволяє мінімізувати участь експерта у вирішенні поставленої задачі. Запропоновано три варіанти злиття даних в задачі побудови карти земного покриття на рівні прийняття рішень, що базуються на використанні точності класифікації кожного класу, маски хмарності зображень та апостеріорної ймовірності пікселів. Реалізацію методів та їх застосування для розв'язання прикладних задач картографування земного покриття розглянуто у п'ятому розділі.

РОЗДІЛ 3

ЕКСПЕРИМЕНТАЛЬНІ ДОСЛІДЖЕННЯ РОЗРОБЛЕНИХ МЕТОДІВ ЗЛИТТЯ СУПУТНИКОВИХ ДАНИХ

В другому розділі даного дисертаційного дослідження були розроблені методи картографування земного покриву для великих територій. Для побудови карт класифікації використовувалися як радарні, так і оптичні супутникові дані, наприклад Landsat, Sentinel-1. В даному розділі буде оцінена ефективність використання часових рядів супутникових даних та розроблені методи злиття даних для побудови карти земного покриву для великих територій. Також будуть описані методи злиття растрових та векторних для покращення карти земного покриву.

3.1 Порівняння кращої мережі та ансамблю

Оскільки вище досліджено, що кращим класифікатором є ансамбль нейронних мереж, саме його будемо використовувати в подальшому. Визначимо оптимальну архітектуру нейромережевого класифікатора. Проведемо експеримент про доцільність використання ансамблю нейронних мереж в порівнянні з окремою нейронною мережею. Основний критерій оцінки класифікатора є точність результатів класифікації на тестовій вибірці.

3.1.1 Дані

Дане дослідження буде проводитись для території Київської області за 2013 рік. В експерименті будемо використовувати дані полігону JEСAM, який включає Київську область [4, 173 – 176].

Площа полігону складає 28,100 км². Основними культурами є кукурудза (25.1% від усіх посівних площ в 2013 р.), озима пшениця (16.1%), соя (12.6%), овочі (10.3%), соняшник (9.3%), ярий ячмінь (6.8%), озимий ріпак (4.0%),

цукровий буряк (1.3%).

Для формування навчальної та незалежної тестової вибірки було виконано наземні дослідження (рис. 3.1). В цілому навчальна і тестова вибірки містять 386 полігонів та включають основні класи земного покриття відповідно до стандартизованої номенклатури класів European Land Use and Cover Area frame Survey (LUCAS) (табл. 3.2). Загальна площа навчальних даних становить 22,700 га.

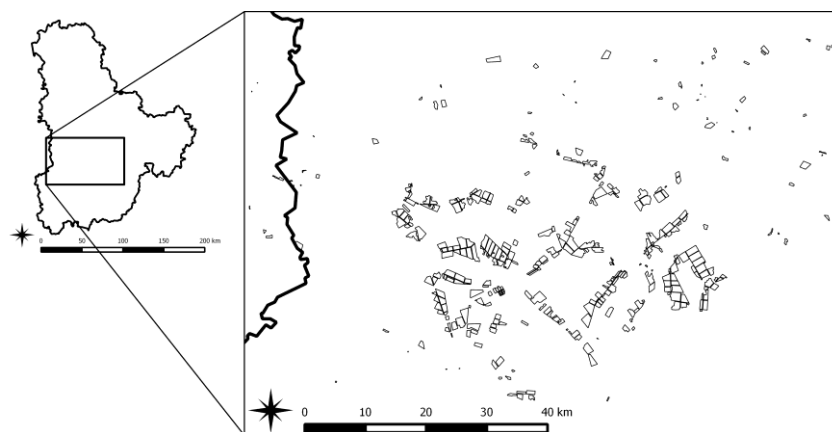


Рис. 3.1 Наземні дослідження на полігоні JESAM

Для проведення експерименту будемо використовувати супутникові дані сенсора Operational Land Imager (OLI) супутника Landsat-8. В якості вхідних ознак класифікатора використовуватимемо 6 спектральних каналів (канали 2–7) з просторовим розрізненням 30 м. Тестовий полігон покривається трьома сценами з координатами 181/24, 181/25 181/26. За вегетаційний період 2013 р. отримано по 6 знімків для кожної із сцен (табл. 3.1).

Таблиця 3.1 –

Супутникові дані по Київській області в 2013 році

Супутник	Дата	Захмареність	Кількість спектральних каналів	Просторове розрізнення (м)
Landsat-8	16.04, 02.05, 18.05, 19.06, 05.07, 06.08 (181/24, 181/25, 181/26)	Менше 30%	30	30

Оскільки в якості вхідних ознак використовується 6 спектральних каналів із 6 знімків, то загальна розмірність вхідного шару класифікатора становить 36. Всі змінні нормуються до діапазону з середнім значенням 0 та середньоквадратичним відхиленням 1. На виході класифікатора маємо 13 класів у відповідності до номенклатури класів LUCAS (табл. 3.2).

Таблиця 3.2 –

Навчальна вибірка відповідно до номенклатури класів LUCAS

№	Код	Клас	Полігон		Площа	
			Кількість	%	Гектари	%
1	A00	Штучні об'єкти	6	1.6	23.0	0.1
2	B11	Озима пшениця	51	13.2	3960.8	17.4
3	B32	Озимий ріпак	12	3.1	937.3	4.1
4	B12, B14	Ярі культури	9	2.3	455.9	2.0
5	B16	Кукурудза	87	22.5	7253.3	31.9
6	B22	Буряк	8	2.1	632.5	2.8
7	B31	Соняшник	30	7.8	2549.0	11.2
8	B33	Соя	60	15.5	3252.3	14.3
9	B19, B39, B40	Інші культури	32	8.3	1364.0	6.0
10	C10, B60	Ліси	17	4.4	1014.3	4.5
11	E01, E02	Необроблювані землі	48	12.4	747.5	3.3
12	F00	Порожня земля	10	2.6	67.2	0.3
13	G01, G02	Водні об'єкти	16	4.1	448.3	2.0
		Разом	386	100	22705.3	100

3.1.2 Аналіз результатів

Для проведення експерименту порівняємо три різні ансамблеві архітектури. В першій з них використано 1 класифікатор, в другій — ансамбль із 5 класифікаторів, та в третій — ансамбль із 7 класифікаторів, які були навчені на спільних даних. Результати експерименту представлені в табл. 3.3. Для порівняння використовувалися наступні метрики: загальна точність класифікації, індекс Каппа, точність «виробника» (ТВ) та точність «користувача» (ТК). Індекс Каппа є статистичною мірою узгодженості між

отриманими класами (по супутниковим даним) та класами, що спостерігаються (наземними дослідженнями). Точність «виробника» показує на скільки добре результат класифікації для даного класу співпадає з тестовими даними. Точність «користувача», в свою чергу, показує на скільки імовірно, що даний клас співпадає з результатами класифікації.

Таблиця 3.3 –

Результати класифікації (ТВ — точність «виробника»; ТК — точність «користувача»)

		Схема 1: Кращий одиничний класифікатор (MLP)		Схема 2: Ансамбль із 5 класифікаторів (MLPs)		Схема 3: Ансамбль із 7 класифікаторів (MLPs)	
	Загальна точність, %	84.60		85.11		85.32	
	Індекс Каппа	0.8144		0.8211		0.8235	
№		ТВ, %	ТК, %	ТВ, %	ТК, %	ТВ, %	ТК, %
1	Штучні об'єкти	74.5	93.2	100.0	97.9	100.0	97.9
2	Озима пшениця	95.6	90.6	96.0	91.9	95.7	91.8
3	Озимий ріпак	94.5	96.1	93.3	99.2	93.5	99.4
4	Ярі культури	12.1	15.3	46.2	38.8	40.6	34.6
5	Кукурудза	92.6	86.6	90.3	86.8	90.5	86.8
6	Буряк	83.0	93.7	94.4	88.0	94.9	89.6
7	Соняшник	86.1	82.1	83.6	84.2	84.1	85.4
8	Соя	66.6	77.1	68.8	76.6	69.7	77.1
9	Інші культури	71.8	76.9	70.2	78.1	70.9	78.0
10	Ліси	96.7	91.9	96.9	91.9	96.9	92.9
11	Необроблювані землі	84.2	88.9	90.7	88.0	91.0	89.0
12	Порожня земля	86.7	88.8	86.7	98.5	86.7	99.0
13	Водні об'єкти	99.3	98.1	100.0	98.0	100.0	98.1

Як видно із табл. 3.3 найвища точність отримана для ансамблю із 7 класифікаторів, які навчались на спільних даних. Матрицю невідповідностей для цього випадку наведено в табл. 3.4.

Таблиця 3.4 –

Матриця невідповідностей для тестової вибірки при застосуванні архітектури 3 (ТВ — точність «виробника»; ТК — точність «користувача»; ПР — помилка першого роду; ДР — помилка другого роду)

Клас	1	2	3	4	5	6	7	8	9	10	11	12	13	ТК,%	ДР,%
1	184	0	0	0	3	0	0	0	0	0	0	1	0	97.9	2.1
2	0	21453	33	51	26	0	29	11	1728	0	2	30	0	91.8	8.2
3	0	1	3788	1	14	0	5	1	2	0	0	0	0	99.4	0.6
4	0	723	0	982	3	0	0	3	1101	0	24	0	0	34.6	65.4
5	0	0	0	177	33972	2	390	4544	34	0	41	0	0	86.8	13.2
6	0	0	0	11	222	3511	118	52	2	0	2	0	0	89.6	10.4
7	0	1	0	10	389	7	13241	1672	91	0	92	0	0	85.4	14.6
8	0	56	0	351	2660	181	1027	14668	29	3	52	0	0	77.1	22.9
9	0	74	230	828	96	0	832	78	7677	0	24	0	0	78.0	22
10	0	27	0	0	78	0	36	11	6	4198	165	0	0	92.9	7.1
11	0	71	1	6	63	0	26	15	159	131	4089	31	0	89.0	11
12	0	0	0	0	0	0	0	0	0	0	4	403	0	99.0	1.0
13	0	0	0	0	0	0	49	0	0	0	0	0	2524	98.1	1.9
ТВ,%	100.0	95.7	93.5	40.6	90.5	94.9	84.1	69.7	70.9	96.9	91.0	86.7	100.0		
ПР,%	0.0	4.3	6.5	59.4	9.5	5.1	15.9	30.3	29.1	3.1	9.0	13.3	0.0		

Результати експерименту показують, що найбільш ефективним є ансамблевий нейромережевий метод. Виявилось, що використання ансамблю нейронних мереж підвищує точність класифікації в порівнянні з окремою нейронною мережею більше ніж на 15% для ярих культур та на 6,5% для необроблюваних земель [3, 4, 177, 178].

При виборі кількості ансамблів, результати експерименти демонструють, що використання одного ансамблю не достатньо, а при використанні п'яти і більше результат практично не змінюється, а лише впливає на час навчання.

Даний метод можна використовувати як для класифікації часового ряду оптичних зображень, так і для комбінації оптичних і радіолокаційних даних.

3.2 Ефективність використання часових рядів супутникових даних в задачі класифікації земної поверхні

Наступний експеримент полягає у визначенні ефективності використання часових рядів супутникових даних для побудови карт земного покриття.

Спочатку буде побудовано карти земного покриття по кожному супутниковому знімку окремо для однієї сцени (результат отримано, виконавши злиття карт по кожному знімку на основі розробленого у другому розділі методу (2.27) на рівні прийняття рішень із урахуванням апостеріорної ймовірності кожного пікселя та маски хмарності). У другій частині даного експерименту, використавши часовий ряд тих самих супутникових знімків, буде побудовано карту земного покриття для тієї ж території. Буде оцінена точність кожної з карт, розроблених першим та другим способом. Для оцінки якості класифікації будемо використовувати показники загальної точності (ОА – overall assigasy) та матриць невідповідностей (формула для визначення загальної точності представлена у другому розділі дисертаційного дослідження (2.20)). В результаті буде визначено ефективність використання часового ряду супутникових даних та кожного знімку окремо, критерієм якої є точність карти земного покриття на незалежній тестовій вибірці.

3.2.1 Дані

Дане дослідження буде проводитись для частини території Київської області за 2010 рік, що покривається одним тайлом (рис. 3.2). Для цього використаємо супутникові дані Landsat-5 з просторовим розрізненням 30 м. Загалом буде використано 4 зображення за такі дати: 08.04.2010, 11.07.2010, 29.07.2010, 30.08.2010. Часовий ряд супутникових даних складається із 24 спектральних каналів (з кожного зображення використаємо по 6 каналів). Хмарність зображень не перевищує 30%.

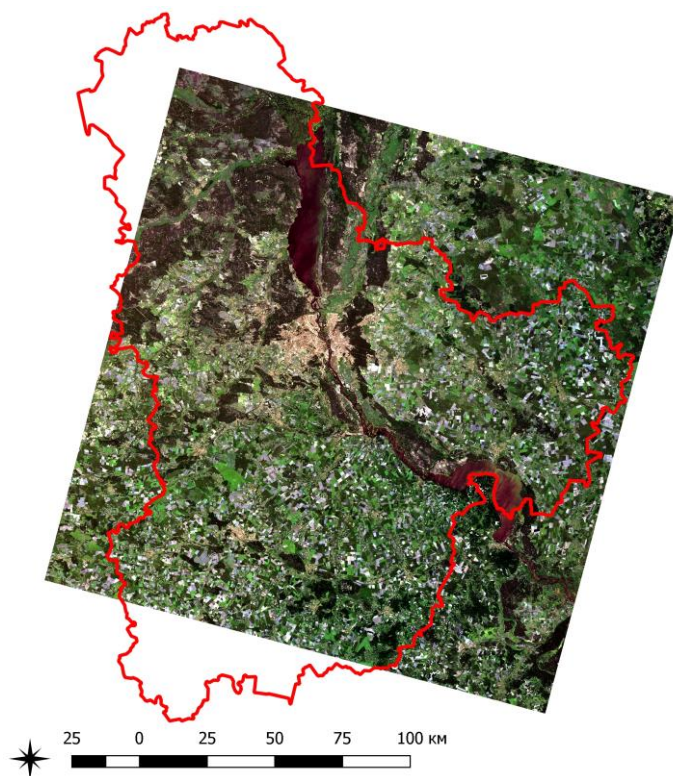


Рис. 3.2 Територія Київської області для якої проводилося експериментальне дослідження (2010 р.)

Класифікацію зображень проведемо за допомогою ансамблю нейронних мереж. Для навчання та перевірки точності використаємо дані, отримані на основі наземних досліджень та фотоінтерпретації (табл. 3.5, рис. 3.3).

Таблиця 3.5 –

Вибірка відповідно до номенклатури класів LUCAS, що використовувалась для побудови карти по Київській області в 2010 році

№	Код	Клас	Полігон		Площа	
			Кількість	%	Гектари	%
1	A00	Штучні об'єкти	0	0	0	0
2	B00	Землі сільськогосподарського призначення	313	30,3	8694	45,7
3	C00	Ліси	199	19,3	2061	10,8
4	E00	Необроблювані землі	371	35,9	5217	27,5
5	G00	Водні об'єкти	150	14,5	3053	16,0
		Разом	1033		19025	

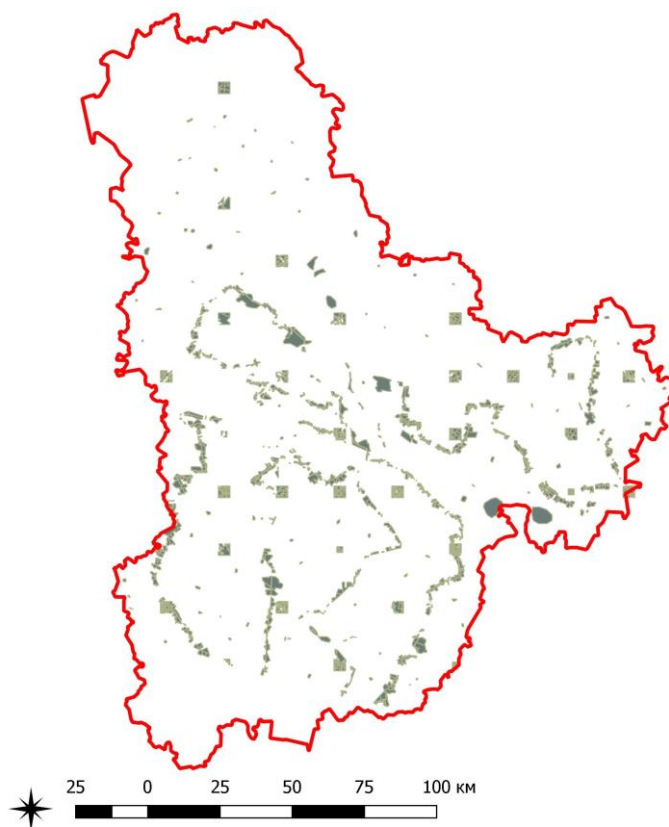


Рис. 3.3 Вибірка по цільовій території (2010 рік)

Для виключення із розгляду штучних об'єктів та населених пунктів буде використана векторна маска (рис. 3.4).

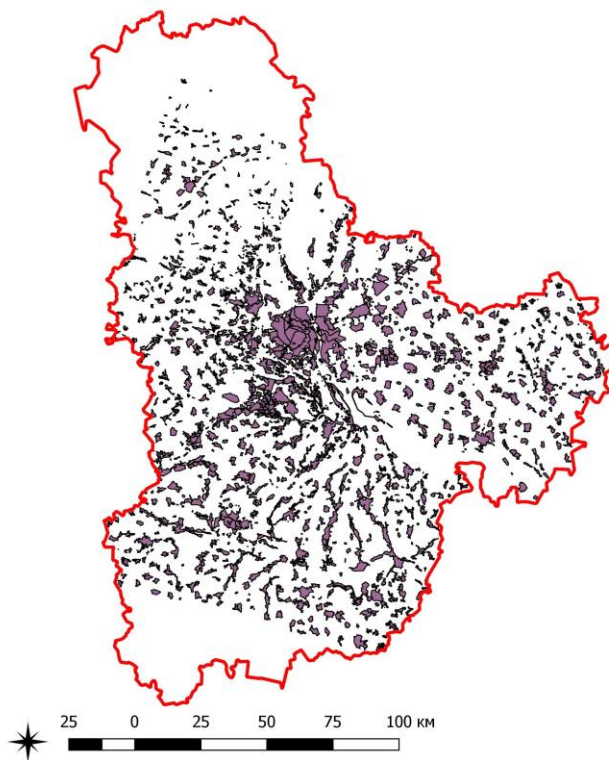


Рис. 3.4 Векторна маска сіл та штучних об'єктів по цільовій території

3.2.2 Аналіз результатів

За допомогою запропонованих підходів побудована карта класифікації для однієї зі сцен Київської області. Для злиття отриманих карт в межах однієї сцени використано метод злиття даних із урахуванням апостеріорної ймовірності кожного пікселя та маски хмарності, описаний у другому розділі (формула 2.27). Отримані результати демонструють перевагу класифікації часового ряду даних, що підтверджується збільшенням точності класифікації на незалежній вибірці більше ніж на 10% для окремих класів (наприклад, необроблювані землі). Детальні точності ТВ та ТК для кожного із класів представлені на рис. 3.5 та в таблиці 3.6.

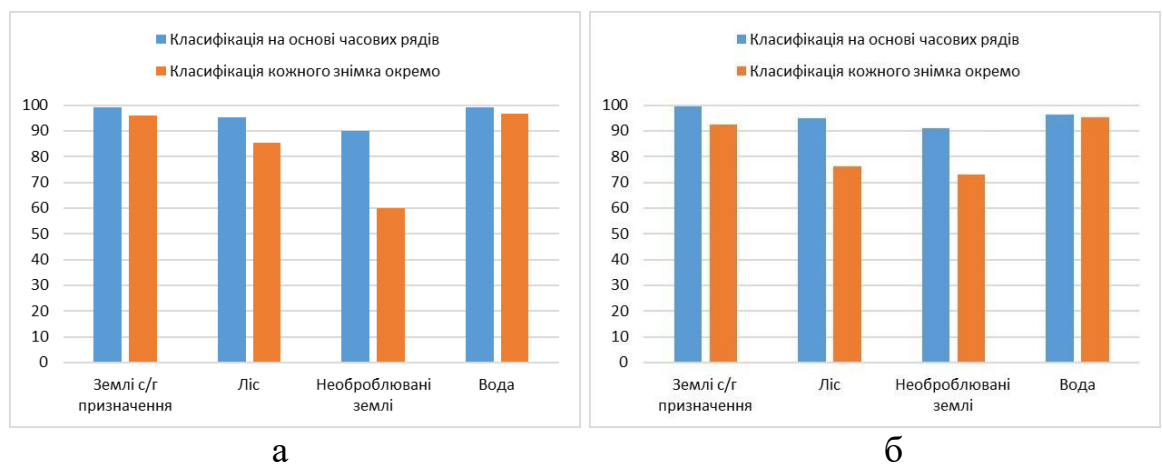


Рис. 3.5 Точність виробника (ТВ) (а), і точність користувача (ТК) (б) для класифікації території Київської області в 2010 році

Проведене експериментальне дослідження, демонструє, що клас водні об'єкти розпізнаються із точністю не нижче 96% в обох випадках. Проте, для класу необроблюваних земель використання часового ряду даних має значну перевагу, ніж при класифікації по окремим знімкам. Це пов'язано із тим, що ідентифікувати даний клас можливо лише за умови аналізу його протягом всього вегетаційного періоду, на відміну від водних об'єктів.

Отримані результати експерименту на незалежній тестовій вибірці підтверджують перевагу класифікації земного покриття на основі часового ряду

супутникових зображень для тих класів, які змінюються під час вегетаційного періоду (наприклад, для земель сільськогосподарського призначення). Для таких класів загальна точність класифікації на основі часових рядів даних на 8% більша, ніж при використанні одного знімку для побудови карти земного покриття.

Таблиця 3.6 –

**Точність виробника та точність користувача по окремим класам
по Київській області в 2010 році**

Клас	Класифікація на основі часових рядів		Класифікації по одному знімку	
	ТВ, %	ТК, %	ТВ, %	ТК, %
Штучні об'єкти	-	-	-	-
Землі сільськогосподарського призначення	98,4	100	96	92
Ліси	95,2	95,1	85,4	77,1
Необроблювані землі	90,3	91,2	60,1	72,2
Водні об'єкти	99	96,7	97	95,5

3.3 Дослідження ефективності злиття оптичних та радарних даних в задачі класифікації сільськогосподарських культур

При виборі супутникових даних для розв'язання задачі класифікації земного покриття в більшості випадків використовують оптичні супутникові зображення. Проте задача часто ускладнюється наявністю пропусків на таких зображеннях, що обумовлено хмарами та тінями. В свою чергу радіолокаційні зображення не залежать від хмарного покриття, але надають інформацію лише на одній довжині хвилі, хоча і для різних комбінацій поляризацій (наприклад,

поляризації для Sentinel-1: HH, VV, VH, HV). На відміну від оптичних зображень, які показують спектральні властивості сільськогосподарських культур, радіолокаційні залежать, в першу чергу, від геометричних властивостей рослин (висота, форма та напрям листків, напрям рядів посівів, тощо).

Таким чином, в даному експерименті буде досліджено ефективність злиття різночасових оптичних та радіолокаційних супутникових зображень для класифікації земного покриву та порівняно результати класифікації з використанням оптичних та радарних даних окремо.

3.3.1 Дані

Тестовою територією для даного експерименту є Київська область за 2015 рік. В якості дистанційної інформації будемо використовувати дані космічних апаратів Landsat-8 та Sentinel-1 (детальна інформація представлена в табл. 3.7), які надаються у відкритому доступі та є безкоштовними.

Таблиця 3.7 –

Супутникові дані по Київській області в 2015 році

Супутник	Дата	Захмареність	Кількість спектральних каналів	Просторове розрізнення (м)
Landsat-8	24.05, 9.06, 25.06, 28.08 (181/24, 181/25, 181/26)	Менше 30%	24	30
Sentinel-1	01.03, 13.03, 25.03, 06.04, 18.04, 30.04, 12.05, 24.05, 05.06, 17.06, 29.06, 11.07, 23.07, 16.08, 28.08	-	30	20
L8+S1	-	-	54	30

Щоб порівняти отримані результати класифікації із незалежними статистичними даними по Київській області необхідно використовувати

супутникові дані, що покривають всю цільову територію. Але для того, щоб повністю покрити супутниковими даними Київську область, необхідно використовувати три тайли Sentinel-1 та Landsat-8 (рис. 3.6).

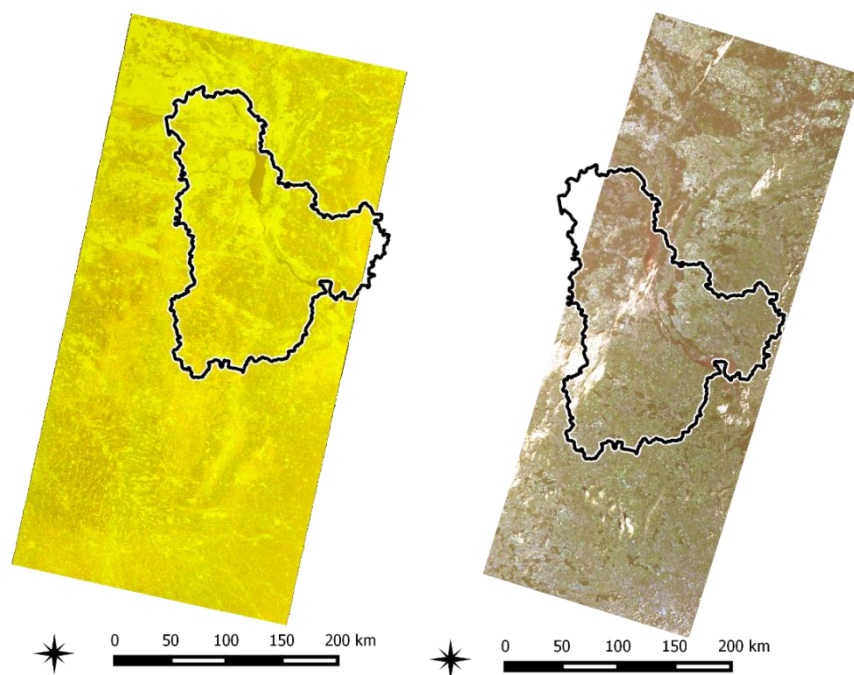


Рис. 3.6 Приклад знімків Sentinel-1 (зліва) та Landsat-8 (справа)

Для збору наземних даних по території Київської області проведено три наземні експедиції (01.05.15, 20.06.15, 12.07.15) (рис. 3.7). Кількісні результати зібраної інформації представлені в табл. 3.8.

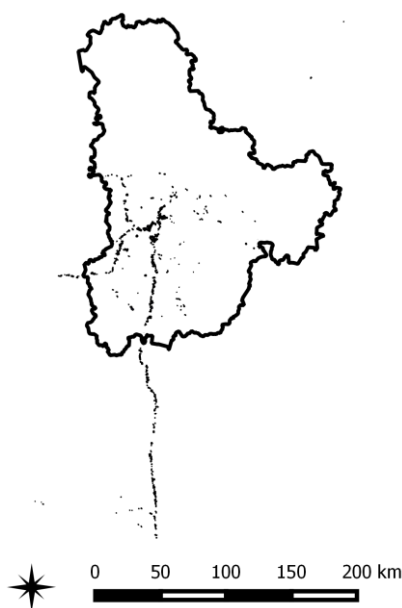


Рис. 3.7 Київська область та зібрана вибірка

Наземні дослідження проводилися по районах, які охоплюють велику кількість полів з різноманітними культурами. Основними культурами є кукурудза, озима пшениця, соя, соняшник, ярий ячмінь, озимий ріпак та цукровий буряк.

В цілому навчальна і тестова вибірки містять 547 полігонів, та включають основні класи земного покриття відповідно до стандартизованої номенклатури класів European Land Use and Cover Area frame Survey (LUCAS) (табл. 3.8) [68]. Загальна площа навчальних даних становить 18966.6 га.

Таблиця 3.8 –

**Навчальна вибірка відповідно до номенклатури класів LUCAS,
що використовувалась для побудови карти по Київській області в
2015 році**

№	Код	Клас	Полігон		Площа	
			Кількість	%	Гектари	%
1	A00	Штучні об'єкти	0	0	0	0
2	B11	Озима пшениця	102	18.6	3695.9	19.5
3	B32	Озимий ріпак	22	4.0	715.9	3.8
4	B12, B14	Ярі культури	11	2.0	296.0	1.6
5	B16	Кукурудза	98	17.9	4329.1	22.8
6	B22	Буряк	8	1.5	860.7	4.5
7	B31	Соняшник	53	9.7	1954.0	10.3
8	B33	Соя	87	15.9	3006.9	15.8
9	B19, B39, B40	Інші культури	0	0	0	0
10	C10, B60	Ліси	49	9.0	2012.3	10.6
11	E01, E02	Необроблювані землі	64	11.7	952.3	5.0
12	F00	Порожня земля	10	1.8	71.4	0.4
13	G01, G02	Водні об'єкти	43	7.9	1072.1	5.7
		Разом	547		18966.6	

3.3.2 Аналіз результатів

Використовуючи запропонований метод злиття різнорідних

супутникових даних в розділі 2 (пункт 2.2 та 2.3), вдалося підвищити як загальну точність класифікації земної поверхні, так і точність окремих класів (табл. 3.9).

Таблиця 3.9 –

Результати класифікації Київської області по різномірних супутникових даних, 2015 р. (ТВ — точність «виробника»; ТК — точність «користувача»)

	Супутник	Landsat-8		Sentinel-1		Landsat-8+ Sentinel-1	
	Клас						
	Загальна точність, %	85,2		91,4		92,7	
	Індекс Каппа	0.83		0.9		0.91	
№		ТВ, %	ТК, %	ТВ, %	ТК, %	ТВ, %	ТК, %
1	Штучні об'єкти	-	-	-	-	-	-
2	Озима пшениця	90,6	94,4	93	95,5	95,2	95,3
3	Озимий ріпак	94,6	78,3	97,6	78,3	97,5	78,4
4	Ярі культури	45,4	76,5	49,8	71,5	50,6	96,7
5	Кукурудза	80,9	84	89,3	89,7	90,6	90,8
6	Буряк	90,4	86,7	94,7	97,3	97,5	98,4
7	Соняшник	94,1	93,1	99,3	93,1	99,2	96,3
8	Соя	63,6	65,6	79,5	83,8	81	83,9
9	Інші культури	-	-	-	-	-	-
10	Ліси	98,7	96,9	99,4	99,2	99,6	99,7
11	Необроблювані землі	79,2	89,3	92,3	85,9	94,5	91,6
12	Порожня земля	64,6	98,7	79,9	69	90,6	82,8
13	Водні об'єкти	97	99,7	98,7	99,9	99,8	99,9

Загальна точність карти класифікації лише за оптичними даними – 85.2%, а додавання радарних даних дозволило збільшити її точність на 7.5%.

Зокрема, додавання радіолокаційних супутникових зображень до оптичних даних значно дозволило покращити точність класифікації літніх культур (наприклад, сої – на 17.4% ТВ та 18,3% ТК; кукурудзи – на 9.7% ТВ та 6,8% ТК). Додавання радарних даних дає можливість здійснювати точнішу класифікацію за рахунок можливості отримання інформації щодо розвитку ходу рослинності протягом всього вегетаційного періоду, чого неможливо

досягти за оптичними даними, оскільки частота зйомка відбувається лише раз на 16 днів і отримані дані можуть бути захмареними. Проте, перевагою оптичних даних є те, що вони дають можливість визначити точні контури полігонів, на відміну від радарних, які містять шум.

На основі отриманих карт розраховані площі посівів для основних сільськогосподарських культур. У табл. 3.10 представлено зіставлення результатів аналізу супутникових даних зі статистичною інформацією за 2015 рік.

Таблиця 3.10 –

Порівняння площ сільськогосподарський посівів отриманих на основі побудованої карти із офіційною статистикою

Клас	Офіційна статистика (1000 га)	Результати класифікації (1000 га)
Озима пшениця	195.2	219.8
Озимий ріпак	35.5	42.8
Кукурудза	241.7	183.2
Цукровий буряк	15.7	13.1
Соняшник	116.9	116.2
Соя	216.5	296.9

Отримана оцінка посівів культур на основі супутникових даних (Landsat-8 і Sentinel-1) за 2015 рік показала, що площі, засіяні озимими культурами, перевищують дані офіційної статистики, зокрема озимої пшениці на 12% та озимого ріпаку на 20%. Щодо інших культур, то отримана площа кукурудзи менша за офіційну площу на 25%, а для сої – перевищує на 37%. Для перевірки достовірності отриманої карти були здійснені наземні дослідження на поля, обрані випадковим чином. Результати наземної експедиції підтверджують достовірність побудови карти земного покриття. Дані офіційної статистики можуть містити похибки, оскільки з різних офіційних джерел площі посіяних культур також не співпадають.

Отже, до описаного експерименту можна зробити висновок, що

використання злиття різнорідних супутникових даних показує кращі результати для побудови карт земного покриття для ідентифікації кожного із класів сільськогосподарських культур та класу необроблюваних земель.

3.4 Експериментальний аналіз розроблених методів злиття результатів класифікації на рівні прийняття рішень

В результаті класифікації з використанням ансамблю багат шарових нейронних мереж перспетронного типу побудовано карти земного покриття для кожної сцени, що покриває територію України, на основі різних супутникових даних (Landsat-4, 5, 7). Ці карти перетинаються між собою (рис. 2.8). Задачею даного експерименту є аналіз трьох методів злиття даних на рівні прийняття рішень (описаних у пункті 2.4 другого розділу дисертаційного дослідження) для побудованих карт земного покриття для великої території, використовуючи карти класифікації по кожному тайлу, та порівняння їх ефективності.

Перший метод — полягає в попиксельному прийнятті рішення на основі ймовірності приналежності даного пікселя певному класу на кожному зображенні, що покриває даний піксель.

Другий метод — рішення про приналежність даного пікселя деякому класу приймається з урахуванням якості знімків.

Третій метод — використовується класифікація на основі часових рядів супутникових даних, а для отримання результуючої карти враховується апостеріорна ймовірність кожного пікселя та маска хмарності.

Також в експерименті потрібно порівняти отримані результати з існуючою картою Globland30-2010 та з офіційною статистикою.

3.4.1 Дані

При розв’язанні задачі побудови карт земного покриття для території України використаємо дані космічних апаратів (КА) Landsat-4, 5, 7 з

просторовою роздільною здатністю 30 м за три роки (1990 р., 2000 р., 2010 р.). Для цього використаємо дані Landsat видимого (Band 1, 2, 3), ближнього інфрачервоного (Band 4, 5) та середнього інфрачервоного (Band 7) спектрів.

Безкоштовні продукти обробки супутникової інформації, використані в даному дослідженні, отримано з геоportалу геологічної служби США (United States Geological Survey) [179].

В дослідженні використано продукт атмосферної корекції даних сенсорів TM та ETM+, створений на основі продуктів рівня обробки L1Gt, або L1T [180], які в свою чергу є радіометрично та геометрично скорегованими даними, що надаються в форматі GeoTIFF. Для перепроєктування супутникових зображень використано бібліотеку GDAL [164]. Приклад результату атмосферної корекції даних Landsat-7 ETM+ зображено на рис. 3.8.

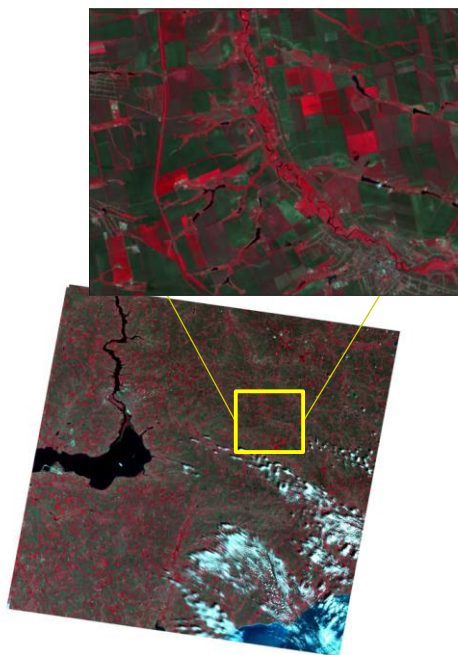


Рис. 3.8 Композит Color InfraRed зображення КА «Landsat-7 ETM+» (10.05.2000) над Запорізькою областю (Path 177 Row 027)

Загалом, для побудови карт земного покриття для 1990, 2000 та 2010 рр. використано 463 супутникові зображення (табл. 3.11). Оскільки для побудови карти земного покриття використовуються ретроспективні супутникові дані, то побудова вибірки буде здійснена на основі фотоінтерпретації, відповідно до

міжнародної номенклатури класів LUCAS [68].

Таблиця 3.11 –

**Супутникова інформація, використана для картографування території
України**

КА	1989 р.	1990 р.	2000 р.	2010 р.	Об'єм файлів, ГБ
«Landsat-4»	23	5	-	-	30,28
«Landsat-5»	-	89	3	145	259,97
«Landsat-7»	-	-	158	40	218,3

До складу побудованих вибірок входять полігони для 6 основних класів (штучні об'єкти, землі сільськогосподарського призначення, ліс, необроблювані землі, відкритий ґрунт, вода), (рис. 3.9). Більш детальна інформація по сформованих вибірках і класах земної поверхні за 1990, 2000 і 2010 рр. представлена в табл. 3.12.

Таблиця 3.12 –

Векторні дані

Клас	Кількість полігонів			Відсоток досліджуваних полігонів від загальної кількості, %			Відсоток досліджуваних полігонів від загальної площі, %		
	2010	2000	1990	2010	2000	1990	2010	2000	1990
Штучні об'єкти	29	14	38	0,22	0,12	0,32	0,04	0,03	0,07
Землі с/г призначення	6278	6000	5942	48,17	49,42	50,96	41,58	40,96	44,49
Ліс	2338	2340	2332	17,94	19,27	20,00	17,28	18,00	20,83
Необроблювані землі	2471	2084	1801	18,96	17,16	15,44	10,14	7,87	7,67
Відкритий ґрунт	315	180	156	2,42	1,48	1,34	1,06	1,47	0,96
Вода	1601	1524	1392	12,29	12,55	11,94	29,90	31,67	25,98
Всього	13032	12142	11661	100	100	100	100	100	100

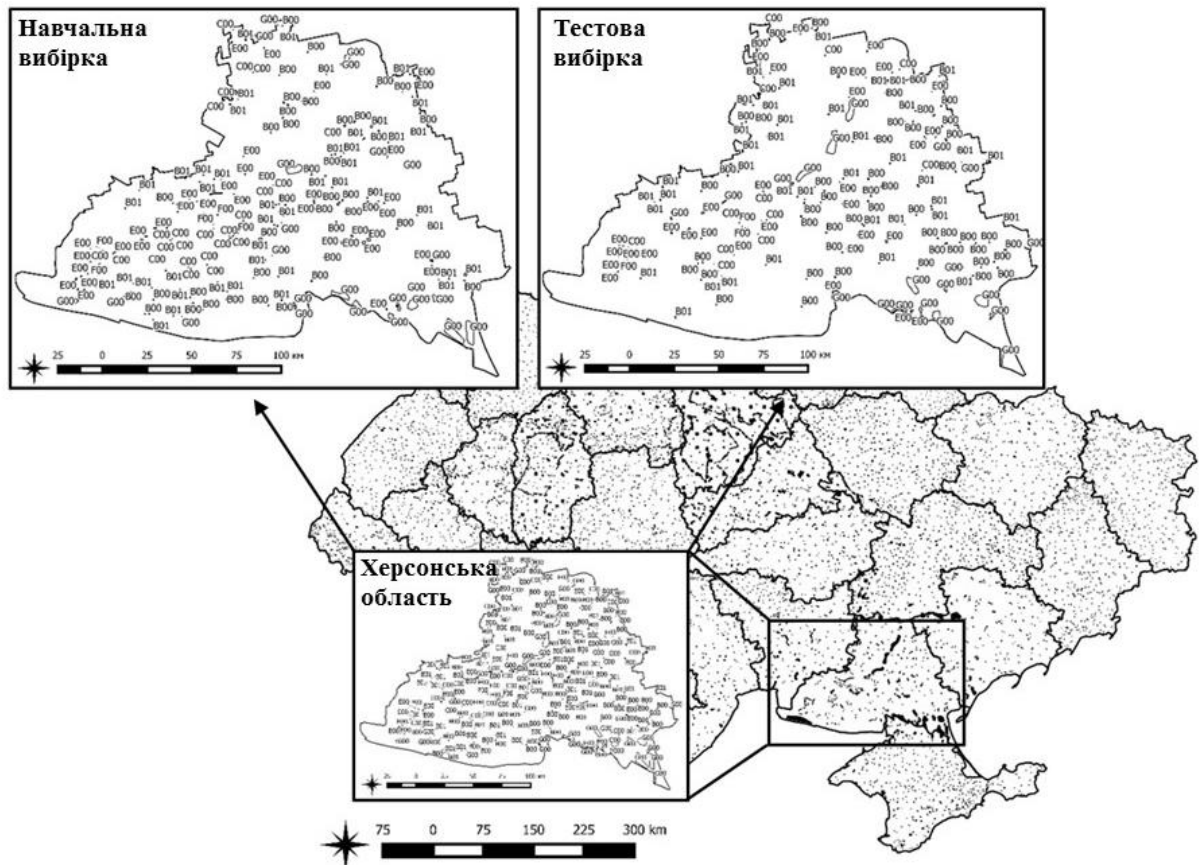


Рис. 3.9 Вибірка для території України за 2010 р. з винесеною окремо вибіркою для Херсонської області

3.4.2 Аналіз результатів

Порівняльний аналіз результатів різних методів злиття даних при побудові карт класифікації показує, що розкид точності «користувача» і «виробника» для основних класів (землі сільськогосподарського призначення, ліс, необроблювані землі і вода) досить великий.

На рис. 3.10-3.11 показані мінімальні, середні та максимальні значення точності виробника і користувача для чотирьох основних класів земного покриття (землі сільськогосподарського призначення, ліс, необроблювані землі і вода) при використанні трьох методів злиття даних. Мінімальні значення для кожного із класів позначені незаштрихованими символами, середні - чорними, максимальні - сірими. На кожному з рисунків вертикальними пунктирними лініями розмежовані результати для кожного із класів. У кожній з секцій, що відповідає певному класу земного покриття, представлені точності, що

забезпечуються трьома різними методами злиття даних. Два перші результати кожного з класів (зліва направо) відповідають методам злиття на рівні прийняття рішень: на основі ймовірності та з урахуванням маски хмарності (класифікація кожного знімка окремо, (2.26, 2.27)). Третій результат у кожній із секцій відповідає злиттю даних із урахуванням апостеріорної ймовірності кожного пікселя та маски хмарності (класифікації на основі часового ряду зображень, (2.28)).

Як видно з рис. 3.10-3.11, використання третього методу злиття даних (класифікації за часовими рядами знімків), забезпечує найкращі результати як для точності «користувача», так і для точності «виробника». Найбільший розкид результатів спостерігається при класифікації необроблюваних земель. Це пояснюється схожістю спектральних характеристик різних типів рослинності (на сільськогосподарських і несільськогосподарських територіях), а значить виникає складність відділення необроблюваних земель від сільськогосподарських культур.

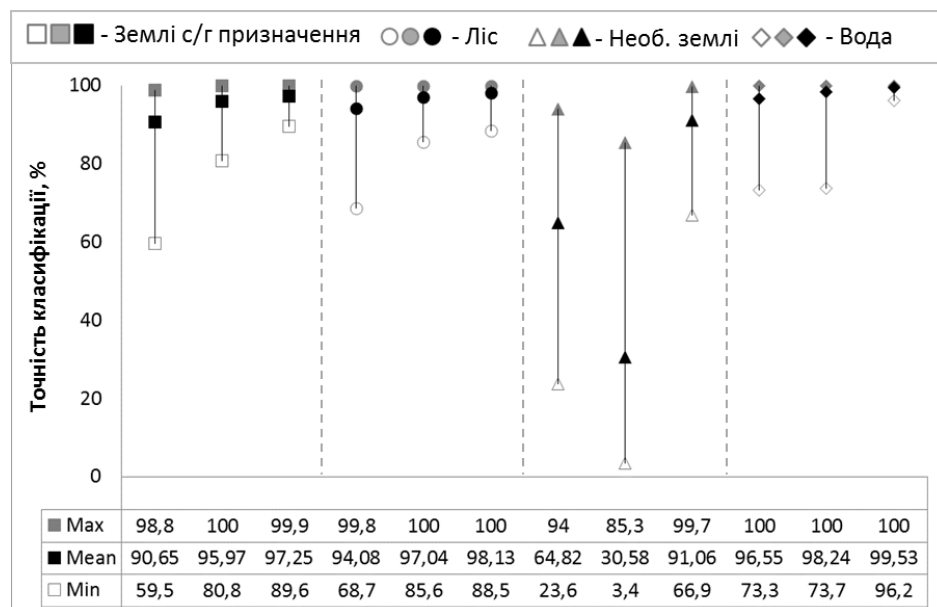


Рис. 3.10 Розкид «точності виробника» для трьох методів злиття даних при побудові карт класифікації за 2010 р. У кожній з секцій (розділених вертикальними пунктирними лініями) зліва направо показані результати злиття на основі ймовірності з урахуванням маски хмарності і злиття даних із урахуванням апостеріорної ймовірності кожного пікселя та маски хмарності (класифікації часового ряду зображень)

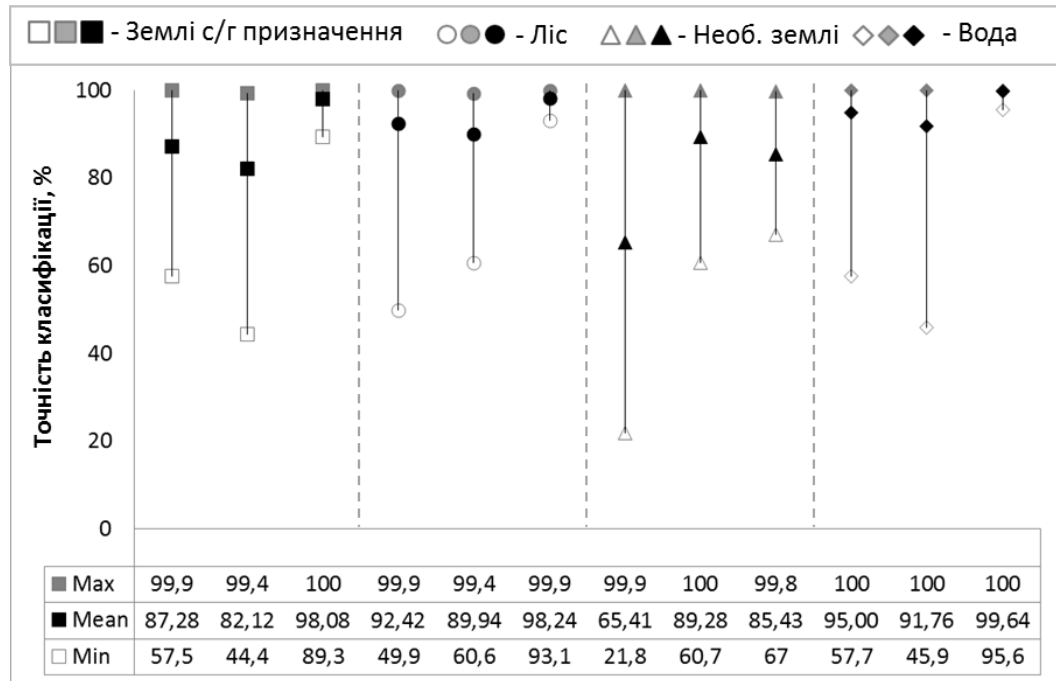


Рис. 3.11 Розкид «точності користувача» для трьох методів злиття даних при побудові карт класифікації за 2010 р. У кожній із секцій, розділених вертикальними пунктирними лініями, зліва направо показані результати злиття на основі ймовірності, з урахуванням маски хмарності і злиття даних із урахуванням апостеріорної ймовірності кожного пікселя та маски хмарності (класифікації часового ряду зображень)

Враховуючи перевагу методу класифікації на основі часових рядів, саме він був використаний для побудови результируючих карт класифікації за 1990, 2000 і 2010 рр. Для кожної з областей загальна точність класифікації становить не менше 88% (в середньому 97%) на незалежній тестовій вибірці. Мінімальна точність класифікації спостерігається в Дніпропетровській (1990) та Львівській (2010 р.) областях. Це пов'язано з недостатньою кількістю супутникових даних в 1990 р. та із зашумленістю (хмарним покривом) більшості знімків, що покривають Львівську область в 2010 р.

Загальна точність класифікації земного покриву для кожної області за 1990, 2000 і 2010 представлена на рис. 3.12.

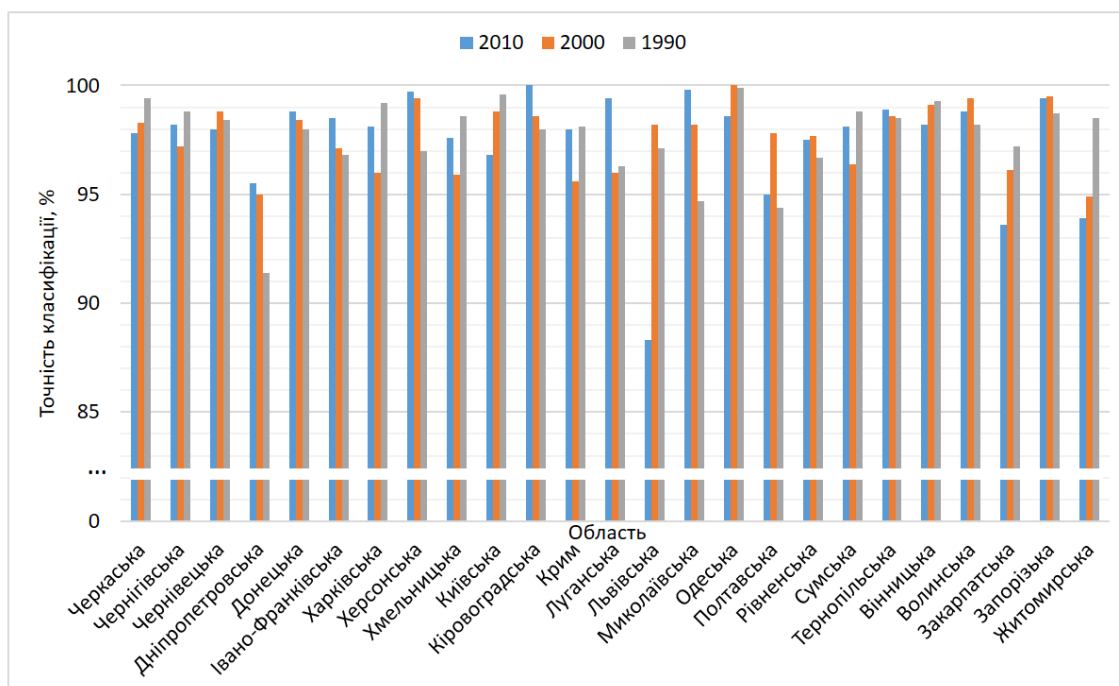


Рис. 3.12 Загальна точність класифікації для кожної області

Отримані карти класифікації земного покриття відрізняються досить високою точністю. Розкид точності «користувача» і «виробника» за основними класами земного покриття представлений на рис. 3.13-3.14.

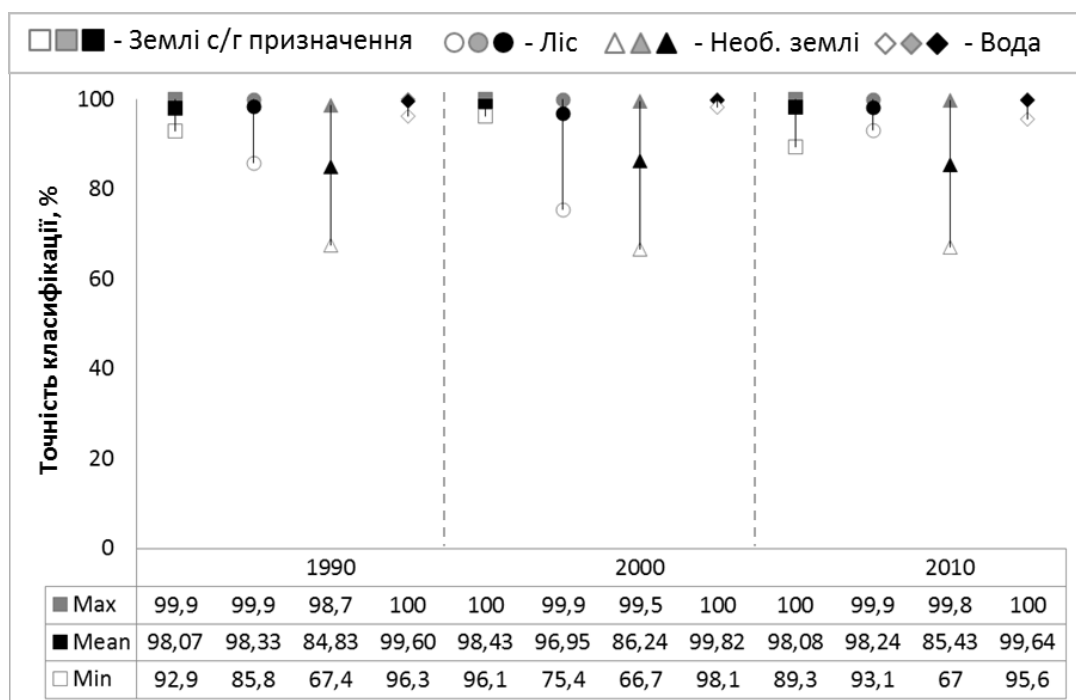


Рис. 3.13 Розкид «точності користувача» для основних класів при класифікації на основі часових рядів зображень (третій спосіб злиття)

Дослідження показали, що найкращі результати класифікації (найменший розкид точності) досягається для класів ліс та вода. Найбільший розкид точності при побудові карти земного покриття припадає на клас необроблюваних земель, мінімальне значення якого складає близько 66%.

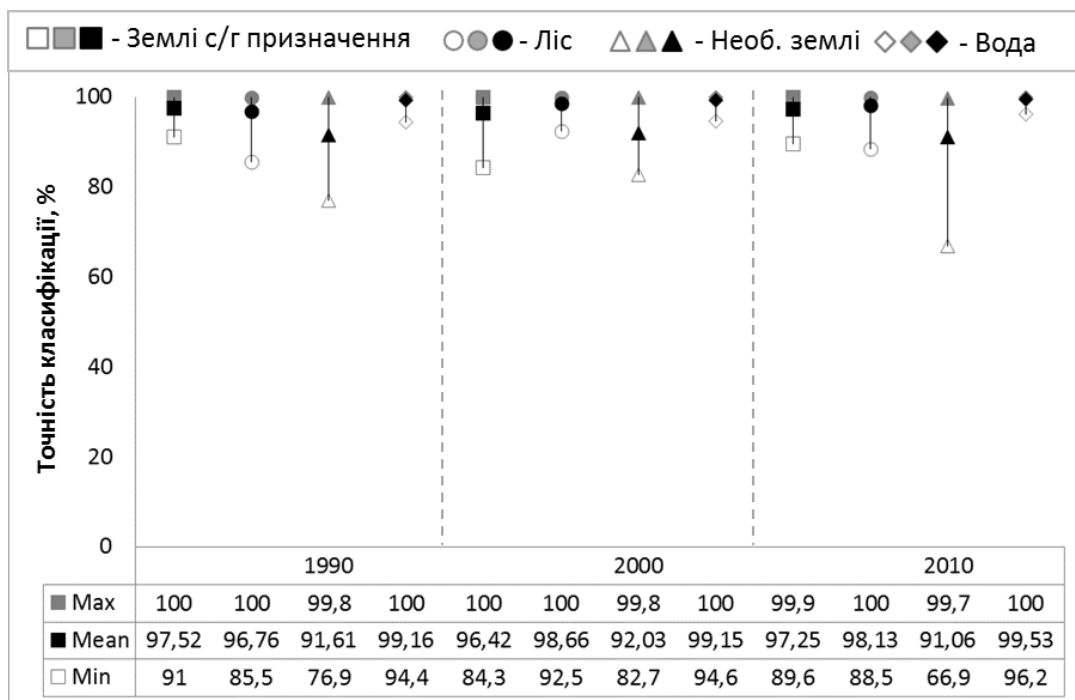


Рис. 3.14 Розкид «точності виробника» для основних класів при класифікації на основі часових рядів зображень (третій спосіб злиття)

Для оцінки точності класифікації використовувалось два підходи: оцінка точності на незалежній тестовій вибірці і порівняння класів областей земельного покриття з офіційною статистикою.

Загальна точність класифікації досягнута в дослідженні була більша 97% (табл. 3.13). Точність для кожного окремого класу не менша 70%. Низька точність була для класу штучні об'єкти (кількість полігонів цього класу у вибірці була невелика та полігони малі за площею) та для необроблюваних земель, оскільки важко відокремити необроблювані землі від земель сільського призначення.

**Загальна точність, точність виробника (ТВ) і точність користувача (ТК)
для класифікації території України в 1990 році, 2000 та 2010 роки та
GlobeLand30-2010**

Рік	2010		2000		1990		GlobeLand30-2010	
Клас	ТВ, %	ТК, %	ТВ, %	ТК, %	ТВ, %	ТК, %	ТВ, %	ТК, %
Штучні землі	100	79.9	73.3	83.9	97.8	92.7	3.4	79.5
Землі сільськогосподарського призначення	97.5	98.5	97.1	98.6	97.5	98.2	85.3	99.4
Ліс	97.2	97.4	98.8	98.4	96.7	98.5	95.9	89.9
Необроблювані землі	90.7	85.4	90.5	84.6	90	82.5	60.5	34.4
Відкритий ґрунт	93.6	96.9	96.2	89.7	94.5	93.4	57.1	0.4
Вода	99.5	99.8	99.5	99.9	99.5	99.7	99.9	96.6
Загальна точність, %	97.5		97.7		97.3		89.7	

В результаті загальна точність для кожної області вище 88%, а середня загальна точність по області приблизно 97%. Інший підхід валідації - порівняння кожного класу з офіційною статистикою – таке порівняння проведено із офіційною статистикою по областях України для кожного періоду часу (1990, 2000 і 2010). Хоча статистичні дані не абсолютно надійні, це єдиний спосіб, щоб оцінити точність класифікації не тільки в конкретній області, а й для всієї території України. Для більшості областей в Україні співвідношення між різницею в офіційної статистики даних і результатами класифікації у діапазоні від 5% до 15% для необроблюваних земель та земель сільського призначення і для класу лісів в межах 5%. Ще одним способом валідації є порівняння точності побудованої карти класифікації для України з глобальною картою рослинного покриву GlobeLand30-2010 (табл. 3.13) з просторовим розрізненням 30 м [181]. Загальна точність результатів класифікації для України на 8% більша, ніж GlobeLand30-2010. Точність класу необроблюваних земель на отриманій карті вища на 30% (точність виробника) та 51% (точність користувача), ніж GlobeLand30-2010. Основна проблема карти GlobeLand30-

2010, що вона не враховує регіональних особливостей.

Отже, аналіз проведених досліджень показав, що метод класифікації на основі часових рядів супутникових даних з використанням апостеріорної ймовірності кожного пікселя та маски хмарності є найкращим для побудови карти земного покриття. Це підтверджується точністю побудованих карт на тестовій незалежній вибірці та порівнянням отриманих карт з даними офіційної статистики та з існуючим продуктом GlobeLand30-2010.

3.5 Експериментальне дослідження ефективності методів покращення карт класифікації за допомогою злиття векторних та растрових даних на рівні прийняття рішень

Оскільки карта класифікації отримана попиксельним методом, то в ній присутні зашумлені скупчення пікселів та неоднорідна структура. В даному експерименті постає задача покращення результатів класифікації для підвищення правдоподібності карти та зменшення її фрагментації. Для цього пропонується використати описані у другому розділі методи злиття даних з урахуванням векторних даних сільськогосподарських посівів та отриманих карт класифікації (пункт 2.4.2 другого розділу).

3.5.1 Дані

Для проведення експерименту використано побудовану карту класифікації для території Київської області в 2013 році (використані дані та метод побудови карти земного покриття для території Київської області в 2013 році наведено в першому експерименті третього розділу). В якості додаткової інформації для покращення отриманої карти використано 16556 векторних меж полігонів для Київської області. Ілюстрація векторних полігонів сільськогосподарських культур для Київської області наведена на рис. 3.15.

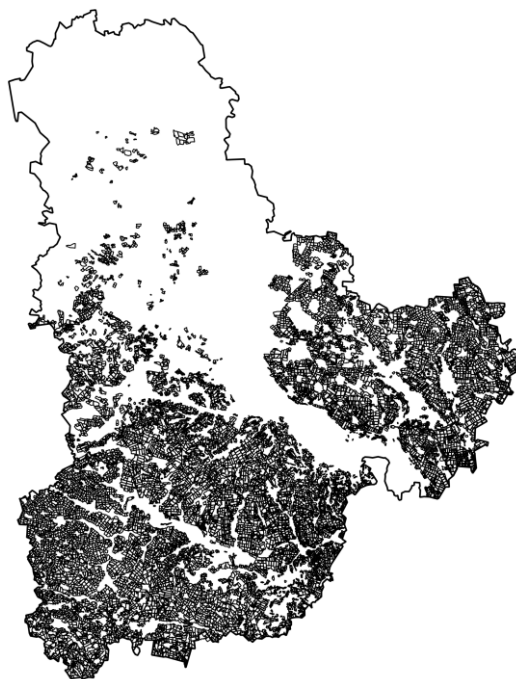


Рис. 3.15 Векторні межі сільськогосподарських культур для Київської області

3.5.2 Аналіз результатів

Для кожного полігону розв'язувалась задача призначення йому єдиного класу, враховуючи, що в межах поля може рости лише одна культура (рис. 3.16) [162].

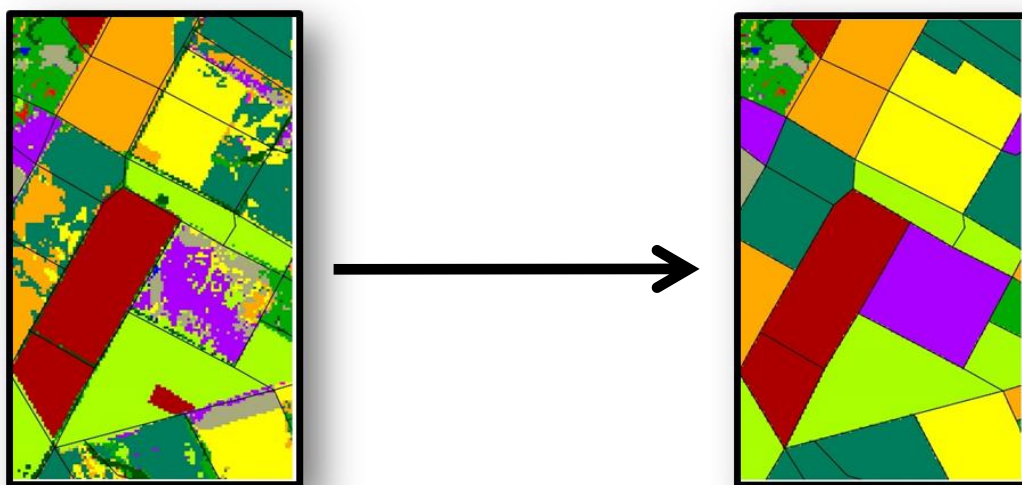


Рис. 3.16 Результат покращення карти класифікації з використанням злиття растрової та векторної інформації Київської області за 2013 р.

На рисунку 3.17 представлений приклад полігону, для якого результат попиксельної класифікації $k_{out}(i, j)$ повністю по формі повторює маску хмарності $m(i, j)$.

При чому захмарених пікселів в цьому полігоні більше ніж не захмарених, в зв'язку з чим, простий метод голосування відносить цей полігон до не правильного класу. В той час як штрафні ваги корегують голосування і назначають правильний результат для полігона.

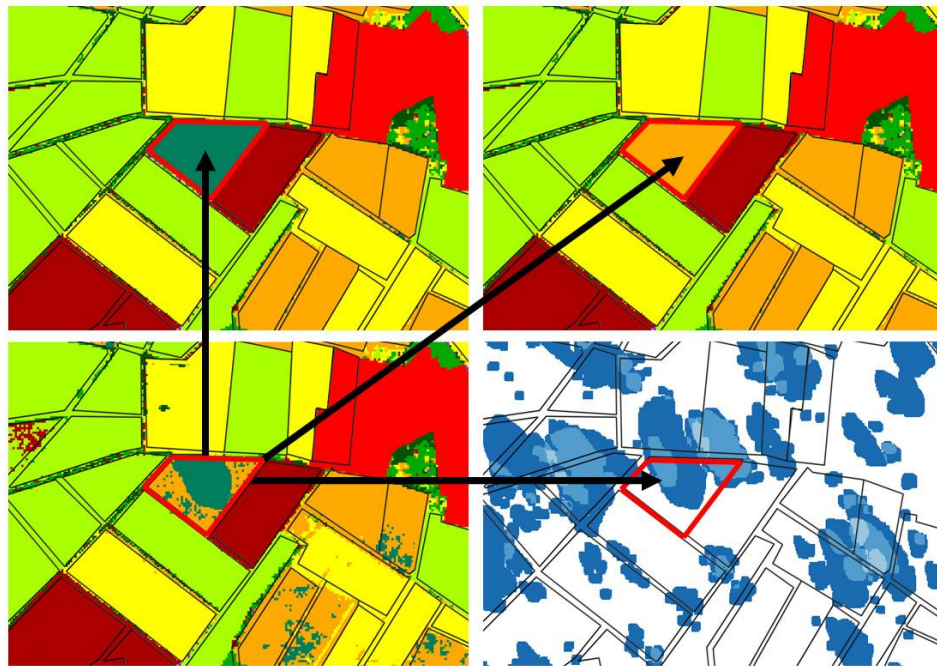


Рис. 3.17 Приклад полігону, для якого метод зваженого голосування (верхній правий) показав кращий результат, ніж звичайний метод голосування (верхній лівий)

Результати для основних сільськогосподарських класів отримані різними методами злиття даних (на рівні прийняття рішень) для карти класифікації Київської області за 2013р. порівняні з офіційною статистикою та з результатами, отриманими на основі початкової карти, представлені в таблиці 3.14.

**Порівняння різних методів злиття даних в межах полігона зі
статистичними даними**

Клас	Офіційна статистика (1000 га)	Попіксельний метод (1000 га / помилка %)	Метод голосування (1000 га / помилка %)	Зважений метод голосування (1000 га / помилка %)
Озима пшениця	187.3	184.5 / - 1.5	168.2 / -10.2	168.7 / -9.9
Озимий ріпак	46.7	59.9 / 28.3	54.1 / 15.8	54.0 / 15.5
Кукурудза	291.7	342.4 / 17.4	338.9 / 16.2	355.3 / 21.8
Цукровий буряк	15.5	11.2 / -27.7	10.8 / -30.4	11.1 / -28.2
Соняшник	108.2	117.6 / 8.7	125.2 / 15.7	125.8 / 16.3
Соя	145.9	168.5 / 15.5	148.3 / 1.6	136.4 / -6.5

Злиття даних за допомогою зваженого методу голосування дозволило покращити карту земного покриття та точніше визначити площі кожного із класів. Іншим методом перевірки ефективності розроблених методів є визначення точності отриманих карт на тестовій незалежній вибірці (табл. 3.15).

Таблиця 3.15 –

**Результати покращення карти класифікації Київської області на основі
методів злиття векторної та растрової інформації, 2013 р. (ТВ — точність
«виробника»; ТК — точність «користувача»)**

	Клас	Попіксельний метод		Метод голосування		Зважений метод голосування	
	Загальна точність, %	83,8		87,7		89,4	
	Індекс Каппа	0.8		0.85		0.87	
№		ТВ, %	ТК, %	ТВ, %	ТК, %	ТВ, %	ТК, %
1	Штучні об'єкти	96,1	64,3	96,1	71,3	96,1	71,3

2	Озима пшениця	95,1	90,7	98,8	94,9	98,8	94,9
3	Озимий ріпак	93,5	97,4	96,5	99,5	96,5	99,5
4	Ярі культури	30,1	28,2	49,8	48,3	49,8	48,3
5	Кукурудза	89,6	84,3	93,2	86,6	93,2	91,6
6	Буряк	93,8	83,5	99,9	99,6	99,9	99,6
7	Соняшник	82,5	86,3	82	91,2	82	91,2
8	Соя	65,9	77,5	73,9	80,2	84,3	82,2
9	Інші культури	68,6	76,6	73,3	79,9	73,3	80
10	Ліси	97,6	89,3	97,5	93,1	97,5	93,1
11	Необроблювані землі	93,1	80,8	93,2	89,5	93,2	89,5
12	Порожня земля	85,6	97,6	86,7	97,6	86,7	97,6
13	Водні об'єкти	99,9	98,6	99,9	98,9	99,9	98,9

Використовуючи запропоновані методи злиття растрової та векторної інформації вдалося збільшити загальну точність класифікації на 5,6%. Зокрема, значно підвищити точність класифікації ярих культур, сої та кукурудзи. Це пов'язано із тим, що відновлені пікселі не завжди є достовірними. Тому при використанні даного підходу (що базується на відновленні пропусків даних) до побудови карти земного покриву метод злиття векторної та растрової інформації на основі зваженого голосування демонструє найкращі результати, свідченням чого є точність отриманої карти класифікації.

3.6 Висновки до розділу

В даному розділі проведено експериментальні дослідження, щодо розроблених методів злиття даних, які зводяться до вирішення задачі класифікації на основі часових рядів даних. Завдяки розробленим методам злиття даних вдалося вирішити задачу класифікації та картографування земного покриву всієї території України за 1990, 2000 і 2010 рік з 30-метровим розрізненням. Загальна точність класифікації по областях становить не менше 88% і в середньому 97% на незалежній тестовій вибірці. Результати отримані при вирішенні задачі картографування, далі можна використовувати для прогнозування врожайності, оцінки ризиків, а також вирішення інших

важливих прикладних задач [182, 183].

В розділі запропоновано використовувати методи покращення карт сільськогосподарських посівів, отриманих на основі супутникових даних, шляхом злиття растрової та векторної інформації, а саме, результатів класифікації супутникових даних та векторних даних про межі полів [6, 7]. Використовуючи векторну інформацію про межі полів (полігонів) запропоновано два методи злиття даних на рівні прийняття рішень. Перший реалізований на основі методу голосування. Конкретному полігону присвоюється значення класу, до якого віднесені більшість пікселів. Другий враховує «надійність» результатів класифікації, яка визначається з урахуванням маски хмарності супутникових зображень, що використовувалися для побудови карти класифікації. Більш надійними вважаються результати класифікації для найменш захмарених пікселів.

РОЗДІЛ 4

АВТОМАТИЗОВАНА ІНФОРМАЦІЙНА ТЕХНОЛОГІЯ КАРТОГРАФУВАННЯ ЗЕМНОГО ПОКРИВУ НА ВЕЛИКИХ ТЕРИТОРІЯХ У МЕЖАХ СЕРВІС-ОРІЄНТОВАНОГО ПІДХОДУ

В даному розділі реалізовані методи злиття даних (що розроблені у другому розділі дисертаційного дослідження) для розв’язання задачі картографування земного покриття для великих територій у формі сервіс-орієнтованої автоматизованої інформаційної технології. Проведено дослідження щодо виявлення вузьких місць сервіс-орієнтованої системи, що впливають на ефективність її роботи.

4.1 Архітектура автоматизованої інформаційної технології картографування земного покриття

Методи для побудови карти земного покриття, що описані в попередніх розділах (другий розділ дисертаційного дослідження), пропонується реалізувати у вигляді автоматизованої інформаційної технології. Реалізувати це у вигляді однієї програми складно, оскільки розроблені методи злиття даних містять процеси, в яких неможливо обійтись без експерта. Тому кожний метод реалізований у вигляді окремого модуля і для отримання карти земного покриття користувач повинен виконувати велику кількість операцій. Це ускладнює та сповільнює процес та при виконанні великої кількості операцій можна допустити помилки. Для полегшення та оптимізації поставленої в дисертаційному дослідженні задачі пропонується реалізувати процес побудови карти земного покриття на основі сервіс-орієнтованого підходу. Перевагою використання сервіс-орієнтованої архітектури є придатність для взаємодії різних і віддалених програмних систем незалежно від платформи. Наприклад, написані модулі програмного коду можуть виконуватись на різних операційних системах (Unix, Windows і тд.). Також сервіс-орієнтована архітектура є

гнучкою для внесення необхідних змін (на відміну від CORBA (Common Object Request Broker Architecture)), де зазвичай використовується локальна мережа з жорстко заданою конфігурацією). Зокрема в задачі картографування земного покриття наведені переваги є важливими, оскільки постійно змінюються вхідні дані (супутникові дані, дані наземних досліджень), що спричиняє зміни в програмних кодах. Тому сервіс-орієнтована архітектура є найбільш оптимальною для розробки автоматизованої інформаційної технології картографування земного покриття, на відміну від існуючих технологій (CORBA, RMI, EJB, Grid і тд.)

В даному дисертаційному дослідженні побудована сервіс-орієнтована архітектура на основі використання Web-сервісу. Розроблений Web-сервіс побудови карт земного покриття із урахуванням відповідності стандартам OGC (Open Geospatial Consortium, <http://www.opengeospatial.org/>), щодо доступу до геопросторової інформації через Інтернет. В загальному схема взаємодії віддаленого клієнта та Web-сервісу представлена на рис. 4.1.

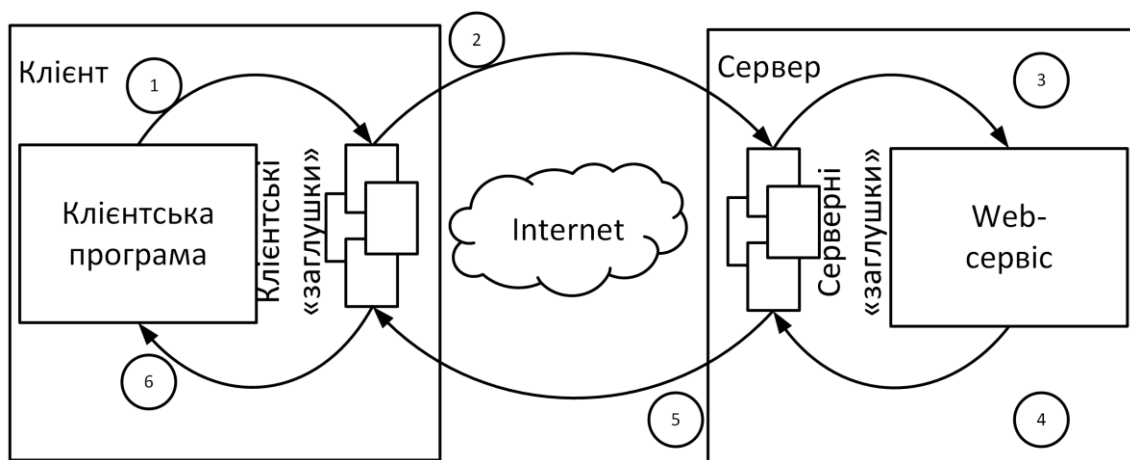


Рис. 4.1 Взаємодія клієнта і Web-сервісу

При використанні Web-сервісу користувачу не потрібно кожного разу вказувати всі параметри підключення до віддаленого вузла. Всі необхідні дії виконуються автоматизовано із залученням так званих stub-класів. Stub-класи - це класи-оболонки, які присутні як у клієнтській, так і в серверній частині розподіленої програми і інкапсулюють реальні дані, якими віддалений клієнт обмінюється з Web-сервісом.

З точки зору реалізації, Web-сервіси - це окремі програмні компоненти, написані на одній з мов програмування, які інкапсулюють певну логіку обробки запитів. Дані, якими віддалений клієнт по мережі обмінюється з Web-сервісом, розміщуються в SOAP-оболонці. Протокол SOAP забезпечує взаємодію з сервером програми, який, в свою чергу, і забезпечує передачу через Інтернет всіх необхідних даних. При цьому в якості транспортного протоколу використовується HTTP, а в ролі засобу передачі виступає Web-сервер.

Розглянемо архітектуру Web-серверу, на якому реалізується автоматизована побудова карт земного покриття. Web-сервер забезпечує розподілений доступ до даних ДЗЗ, їх обробку та надання відповідних результатів. Для роботи запропонованої автоматизованої інформаційної технології дані можуть надходити як із зовнішніх джерел на постійній основі (дані із супутників, результати їх обробки, дані наземних досліджень, векторні дані), так і за запитом користувачів системи, коли її ресурси використовуються для проходження одного з етапів обробки інформації. Зберігання супутникових даних реалізується у вигляді багаторівневого архіву, де на різних рівнях знаходиться інформація, для якої характерна різна частота використання. Поділ даних реалізований за трьома рівнями, представлений на рис. 4.2.

На рівні зберігання даних використовуються наступні архівні компоненти:

1. Оперативний архів, де зберігається нова інформація, яка надходить із різномірних джерел. Для зберігання такої інформації зазвичай використовуються жорсткі диски, які забезпечують мінімальний час доступу до даних.

2. Короткостроковий архів, де зберігається інформація за деякий короткий проміжок часу (тиждень-місяць). Для її зберігання використовують накопичувачі з відносно невеликим часом доступу до інформації (стрічкові накопичувачі, бібліотеки).

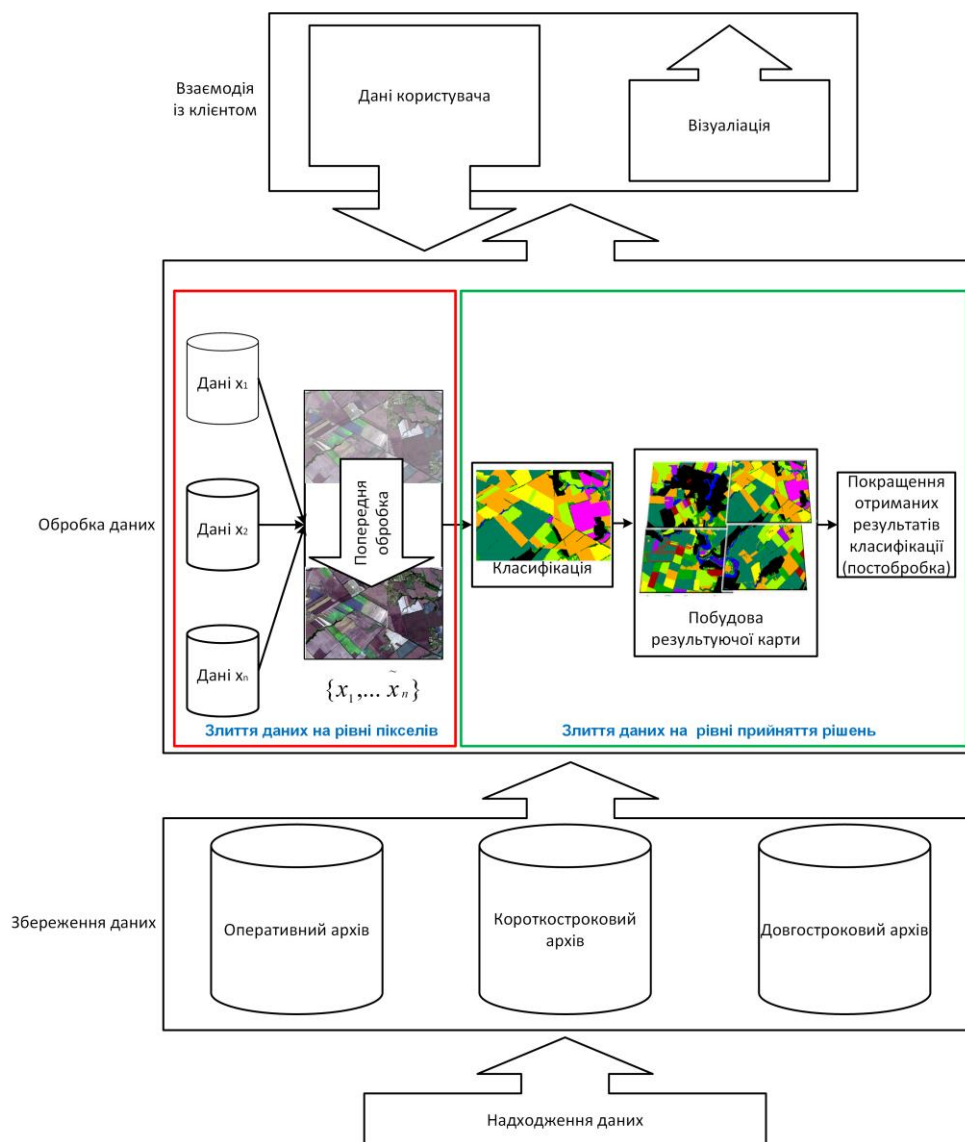


Рис. 4.2 Високорівнева архітектура автоматизованої інформаційної технології для побудови карти земного покриття

3. Довгостроковий архів, в якому зберігається інформація за великий проміжок часу (роки). У деяких випадках такий архів може бути неавтоматизованим, в інших - використовується для зберігання даних у форматах з високим ступенем стиснення. Час доступу до даних на цьому рівні може становити кілька годин або навіть діб в залежності від ступеню стиснення архіву.

На кожному рівні можливе розділення на підрівні, проте загальна схема розподілу даних відповідно до їх необхідності, яка виражається в частоті

запитів до цих даних, залишається незмінною. Підсистема зберігання даних надає програмний інтерфейс для доступу до об'єктів зберігання.

Обробка інформації, що надійшла з архіву або від користувача, проходить три етапи. На першому етапі проводиться попередня обробка зображень, перетворення до єдиного формату, приведення даних до однакових проекцій, однакового просторового розрізнення пікселів зображення. Також з використанням методу злиття даних на рівні пікселів готується вектор інформативних ознак для подальшої класифікації, на чому і завершується етап обробки супутникових даних (як описується в розділі 2, пункт 2.2). Наступний етап обробки інформації включає в себе класифікацію супутникових даних, постобробку результатів (злиття на рівні прийняття рішень) та покращення отриманої карти класифікації на основі злиття растрової та векторної інформації (описано в розділі 2, пункт 2.4). Останній етап обробки даних призначений для приведення отриманого результату до зручної візуалізації і може включати в себе зміну формату даних, перепроєктування, вибір палітри та ін.

Таким чином дані підготовлені до останнього рівня (взаємодія із клієнтом), основною функцією якого є представлення результатів побудови карт земного покриття та подальшої оцінки результату, порівняння із статистичними даними і тд. Інформація, що надається, візуалізується в звичайних форматах для графічних файлів (JPEG, PNG, TIFF), що підтримується більшістю програмних систем. У тих випадках, коли доступ до інформації здійснюється за допомогою протоколів Інтернет, візуалізація виконується з використанням спеціальних програмних засобів, що відносяться до класу картографічних серверів (MapServer). Крім візуальної інформації, що має відносно невеликий обсяг, користувач має програмний інтерфейс доступу до результатів обробки. Такий інтерфейс надається через підсистему зберігання даних, куди надходять результати обробки. Тому, важливою задачею при побудові сервіс-орієнтованої системи (СОС) є дослідження її продуктивності та виявлення пороблених місць, що найбільше на неї впливають.

4.2 Структурна модель сервіс-орієнтованої системи в задачі картографування земного покриву

Для подальшого аналізу та оцінки ефективності сервіс-орієнтованої системи введемо її структурну модель, яка основана на використанні стандартів і добре документованих сервісних інтерфейсів [184]. Таку модель можна представити у вигляді схеми (рис. 4.3). Оскільки в багатьох випадках на ефективність роботи інформаційної системи впливають саме апаратні ресурси, то на даному рисунку увага надається саме їм.

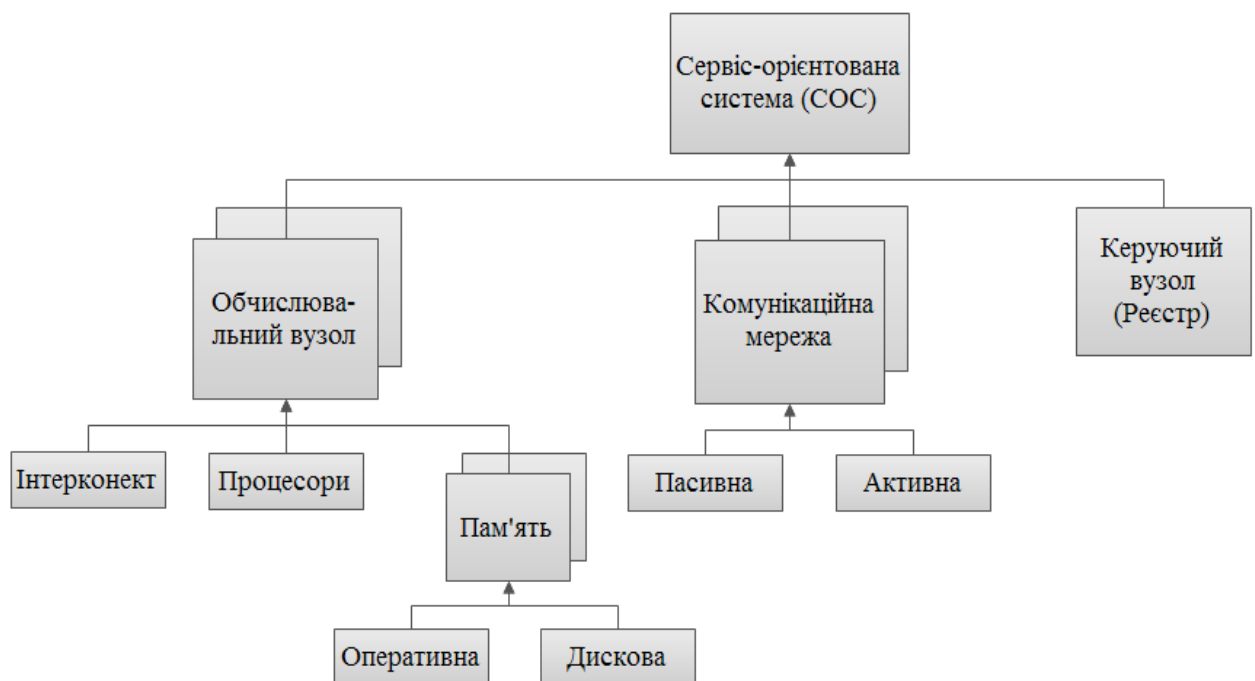


Рис. 4.3 Структурна модель СОС

Головними компонентами СОС є:

- обчислювальний вузол – це багатопроцесорний комп'ютер, на якому виконуються сервіси. Завдання користувача може займати і декілька обчислювальних вузлів;
- комунікаційна мережа – забезпечує обмін даними між компонентами СОС, її користувачами і постачальниками сервісів. Її головне завдання – передача даних без помилок і спотворення [134, 184];
- керуючий вузол – підтримує задачі, які необхідно розподілити між

обчислювальними вузлами, а також каталог сервісів (реєстр) доступних в СОС, з їх фізичним місцем розташування, версією і терміном дії сервісів. Реєстр сервісів є одним з основних компонентів архітектури СОС.

Для вирішення задачі побудови карти земного покриття для великої території важливою є оцінка продуктивності окремих елементів розподіленої системи. Для цього можна скористатися структурно-функціональним аналізом [185, 187].

4.2.1 Постановка задачі структурно-функціонального аналізу сервіс-орієнтованих систем

Враховуючи високу вартість створення розподілених систем, в яких задіяна високопродуктивна техніка, до їх розробки потрібно підходити з використанням системного аналізу, зокрема методу структурно-функціонального аналізу.

Розглянемо загальний випадок структурно-функціонального аналізу сервіс-орієнтованої системи.

Кожен q -й рівень ієрархії системи складається з P_q функціональних елементів. Тоді множина всіх функціональних елементів системи можна описати таким чином:

$$V = \{V_{qp} \mid q = \overline{1, N}, p = \overline{1, P_q}\}, \quad (4.1)$$

де N — загальна кількість рівнів ієрархії; V_{qp} — p -й функціональний елемент (ФЕ) q -ого ієрархічного рівня.

Кожен ФЕ V_{qp} системи характеризується вектором показників

$$x_{qp} = (x_{qpj} \mid j = \overline{1, n_{qp}}), \quad (4.2)$$

де n_{qp} — кількість показників ФЕ V_{qp} , що виконують набір функцій

$$\Phi_{qp} = (f_{qpj} \mid j = \overline{1, m_{qp}}). \quad (4.3)$$

Кожна з функцій f_{qpj} в (4.3) залежить від значень вектора x_{qp}

$$f_{qpj} = f_{qpj}(x_{qp}) \quad (4.4)$$

і впливає на реалізацію вимог, що пред'являються до ФЕ і системи в цілому.

Згідно [186] вигляд функцій (4.3), (4.4) визначається в процесі системного аналізу.

Нехай

$$X = \{x_{qp}, p = \overline{1, P}, q = \overline{1, N}\} \quad (4.5)$$

узагальнена множина показників всіх ФЕ. Для простоти запису відмовимося від потрібної індексації. Будемо вважати, що загальна кількість показників всіх ієрархічних рівнів системи рівна N_0 . Тоді узагальнений вектор показників можна представити у вигляді

$$x = (x_1, x_2, \dots, x_{N_0})^T, x \in R^{N_0}. \quad (4.6)$$

Одним з головних завдань структурно-функціонального аналізу є визначення перетворення

$$F : X \rightarrow Y, \quad (4.7)$$

що діє з множини X допустимих показників системи в простір Y необхідних властивостей по набору кількісних вимог [187].

При цьому необхідно розв'язати задачу, що дозволяє одночасно визначити структуру сервіс-орієнтованої системи в цілому, структуру функціональних елементів всіх ієрархічних рівнів і вид перетворення (4.7).

Проаналізуємо елемент СОС, що включає сховище даних, тобто файловий архів [188]. Для цього досліджується поведінка сервіс-орієнтованої системи у випадках, коли для вирішення завдання і виконання обчислень необхідно завантажити зі сховища один або декілька файлів.

4.2.2 Структурно-функціональний аналіз сервіс-орієнтованої системи з урахуванням типів вирішуваних завдань

Розглянемо приклад задачі структурно-функціонального аналізу для сервісного вузла, що вирішує в СОС проблеми певних типів. Введемо множину типів програмного коду K , а також індикатор типу для коду: якщо d — код (послідовність інструкцій), то $type(d)$ — його тип. Задача складається з частин

програмного коду різного типу. Експериментально можна побудувати функцію ефективності процесора при обробці коду певного типу $w_1: K \times Arch \rightarrow R$, де $Arch$ — архітектура процесора, а R — множина дійсних чисел. Існують організації, які надають таблиці значень таких функцій для еталонних задач, які стали загальноприйнятими стандартами для оцінки продуктивності комп'ютерних систем. Наприклад, організація виробників комп'ютерів SPEC (System Performance Evaluation Corporation) [187] публікує результати досліджень продуктивності процесорів, мультипроцесорних систем, файлових серверів і т.д. При цьому в якості еталонних задач можуть використовуватися програми, що виконують тільки базові операції (подібні додаванню і множенню, наприклад програма Dhrystone), модельні задачі (наприклад, решето Ератосфена і Ханойські вежі), ядра (наприклад, програма Linpack) або реальні прикладні задачі (наприклад, модель чисельного прогнозування погоди WRF) [188].

Тоді з урахуванням продуктивності процесора з частотою f_r , при завантаженні z для коду типу k та архітектури $arch \in Arch$

$$f_{cpu} = w_1(k, arch) f_r (1 - z). \quad (4.8)$$

Розглянемо вузол, що містить N процесорів. Нехай продуктивність інтернет контенту (кількість інформації, переданої між процесорами за одиницю часу) дорівнює f_{int} . Оскільки оброблену процесором інформацію необхідно передавати іншим процесорам, при використанні N процесорів продуктивність вузла не збільшиться рівно в N разів. Нехай $w_2: K \times Arch \rightarrow R$ — функція, що визначає частину даних, які необхідно передати між процесорами під час обчислень при виконанні коду одиничної (еталонної) довжини типу K . Таблицю значень цієї функції можна побудувати для кожної задачі, що розв'язується в системі. Тоді при обробці об'єму інформації C необхідно передати інформацію об'ємом $C w_2(k, arch) = C\beta$. Тоді продуктивність вузла (рівна кількості оброблених даних) обчислюється за більш складною формулою

$$f_{node}(x_{node}) = \begin{cases} f_{proc}(x_{proc}) f_{mem}(x_{mem}) f_{con}(x_{con}) n_{proc}, & \text{для об. вузла} \\ f_{proc}(x_{proc}) f_{stor}(x_{stor}) f_{con}(x_{con}) n_{proc}, & \text{для сховища} \end{cases}$$

Кількість інформації C спочатку буде оброблено процесором за час C / f_{cpu} , потім ця інформація передається іншому процесору за час $C\beta / f_{int}$. Загальний час буде складати $C / f_{cpu} + C\beta / f_{int}$. Тоді продуктивність вузла з N процесорами буде становити

$$f_{N-proc} = \frac{CN}{C / f_{cpu} + C\beta / f_{int}} = \left(\frac{1}{f_{cpu}} + \frac{\beta}{f_{int}} \right)^{-1} N. \quad (4.9)$$

Коефіцієнт необхідної оперативної пам'яті $m_1 = m_1(K)$ показує, скільки оперативної пам'яті необхідно використовувати для обчислення задачі типу K одиничної довжини. Тоді сумарний обсяг необхідної оперативної пам'яті, необхідний для розв'язання задачі, дорівнює $S = \max_{1 \leq i \leq l} (m_1(\text{type}(d_i)) \text{size}(d_i))$, де d_i - i -та послідовність інструкцій коду, $\text{size}(d_i)$ - розмір коду d_i . Якщо такий обсяг вільної пам'яті відсутній, необхідно використовувати віртуальну пам'ять, тобто пам'ять жорсткого диска. Тому введемо коефіцієнт необхідної пам'яті жорсткого диска $m_2 = m_2(K)$.

Продуктивність оперативної пам'яті вважається постійною $f_{RAM} = \text{const}$. Для обчислення продуктивності жорсткого диска (віртуальної пам'яті) використовуємо наступний підхід. Відомо, що продуктивність жорстких дисків істотно залежить від порядку зчитування з нього інформації. У разі послідовного зчитування продуктивність максимальна, а в разі випадкового доступу вона може зменшитися в десятки разів. Очевидно, що розташування даних на диску і порядок зчитування залежать від коду. Таким чином, продуктивність жорсткого диска являє собою функцію типу коду, і корисна продуктивність пам'яті визначається як $f_{HDD} = f_{HDD}(x_{prod}, k)$ — кількість даних, які можуть бути записані або зчитані з жорсткого диска, враховуючи особливості цих запитів і характеристики конкретного диска [187].

Обчислимо корисну продуктивність вузла. Враховуючи, що час витрачений на обчислення, підсумовується з часом, витраченим на читання або

запис, одержуємо, що корисна продуктивність при обробці коду типу K (сума зворотних величин) визначається виразом

$$f_{node} = \left(\frac{1}{f_{N_proc}} + \frac{m_1 \alpha_{RAM}^1 + m_2 \alpha_{RAM}^2}{f_{RAM}} + \frac{m_1 \alpha_{HDD}^1 + m_2 \alpha_{HDD}^2}{f_{HDD}} \right)^{-1}, \quad (4.10)$$

де коефіцієнти $\alpha_{RAM}^{1,2}$ та $\alpha_{HDD}^{1,2}$ показують, як буде використовуватися пам'ять. Коефіцієнт α_{HDD}^1 визначає, яка частина даних, що повинні зберігатися в оперативній пам'яті, записується на диск, α_{HDD}^2 — яка частина даних буде зберігатися на жорсткому диску (як і заплановано), α_{RAM}^1 — яка частина даних буде зберігатися в оперативній пам'яті (як і заплановано), α_{RAM}^2 — яка частина даних, призначених для зберігання на жорсткому диску, буде зберігатися в оперативній пам'яті (при використанні кешування). Таким чином, враховується можливість використання віртуальної пам'яті і кешування. При цьому $\alpha_{RAM}^1 + \alpha_{HDD}^1 = 1$ та $\alpha_{RAM}^2 + \alpha_{HDD}^2 = 1$.

4.3 Результати чисельного моделювання та оцінка ефективності передачі даних в сервіс-орієнтованій системі

Експеримент 1.

Оцінка ефективності передачі одного файлу даних в сервіс-орієнтованій системі

Проілюструємо застосування запропонованого підходу для оцінки продуктивності сервіс-орієнтованої системи при розв'язанні задач різного типу.

Загальна постановка задачі зводиться до наступного.

Задана загальна архітектура обчислювального вузла (рис. 4.2) і типи виконуваних завдань. Потрібно на основі запропонованого підходу до структурно-функціонального моделювання СОС визначити наступні характеристики:

- 1) продуктивність СОС при виконанні однотипних задач;
- 2) виявити слабкі (критичні) місця СОС (тобто ті місця, які обмежують

сумарну продуктивність системи).

Розглянемо систему, яка має такі властивості:

1. Система, що складається з 3 однопроцесорних комп'ютерів і 2 кластерів, при цьому один з однопроцесорних комп'ютерів використовується в якості керуючого вузла, що забезпечує тільки надходження задач в сегмент користувача.

2. Взаємодія між вузлами здійснюється за допомогою 2 маршрутизаторів.

3. Пропускна здатність мережі становить $1000 \text{ Мбіт} / \text{с} = 125 \text{ Мб} / \text{с}$.

4. Для простоти будемо вважати, що продуктивність процесорів у формулі (4.8) пропорційна тільки тактовій частоті f_r і не залежить від типу виконуваного коду (код вважаємо однорідним).

5. Будемо вважати, що в розглянутому прикладі однопроцесорні комп'ютери мають частоту $f_r = 1 \text{ ГГц}$.

6. Оперативна пам'ять однопроцесорного комп'ютера складає 256 Мб.

7. Жорсткий диск такого комп'ютера має розмір 60 Гб.

8. Кожен обчислювальний вузол містить 16 процесорів з частотою 2,5 ГГц. Обсяг локальної оперативної пам'яті обчислювального вузла становить 128 Мб на кожен процесор і ємність кожного з 16 жорстких дисків – 80 Гб.

Припустимо, що для вирішення кожного завдання перед виконанням обчислень необхідно передати зі сховища один файл розміром в 500 Мб.

Потрібно. Оцінити продуктивність сервіс-орієнтованої системи.

Для вирішення даної проблеми спочатку обчислимо продуктивність вузла. Враховуючи, що час, витрачений на обчислення, підсумовується з часом, витраченим на читання або запис, то корисна продуктивність визначається виразом [187] при наявності необхідного обсягу даних, використовуючи формулу 4.9, обчислимо продуктивність кластера, що в цілому складає:

$$f_{node} = \left(\frac{1}{8,33 \cdot 10^9} + \frac{0,1 \cdot 1}{10^9} \right)^{-1} 16 = 7,27 \cdot 10^{10} \text{ байт} / \text{с}.$$

Для персонального комп'ютера ця величина складає:

$$f_{node} = \left(\frac{1}{4 \cdot 10^9} + \frac{0,2 \cdot 1 + 0 \cdot \alpha_{RAM}^2}{10^{10}} + 0 \right)^{-1} = (2,5 \cdot 10^{-10} + 2 \cdot 10^{-11})^{-1} \approx 3,70 \cdot 10^9 \text{ байт/с.}$$

Нехай в середньому на персональні комп'ютери кожену секунду подається x_1 та x_2 задач. На кластери відповідно подається y_1 та y_2 задач за секунду. Тоді для забезпечення виконання цих завдань на маршрутизатор, показаний на рис. 4.4 зліва, кожену секунду потрібно завантажувати зі сховища $5 \cdot 10^8(x_1 + x_2)$ байт, а на маршрутизатор праворуч — $5 \cdot 10^8(y_1 + y_2)$ байт. Крім завантаження даних для вузлів, необхідно завантажувати завдання безпосередньо на обчислювальні елементи.

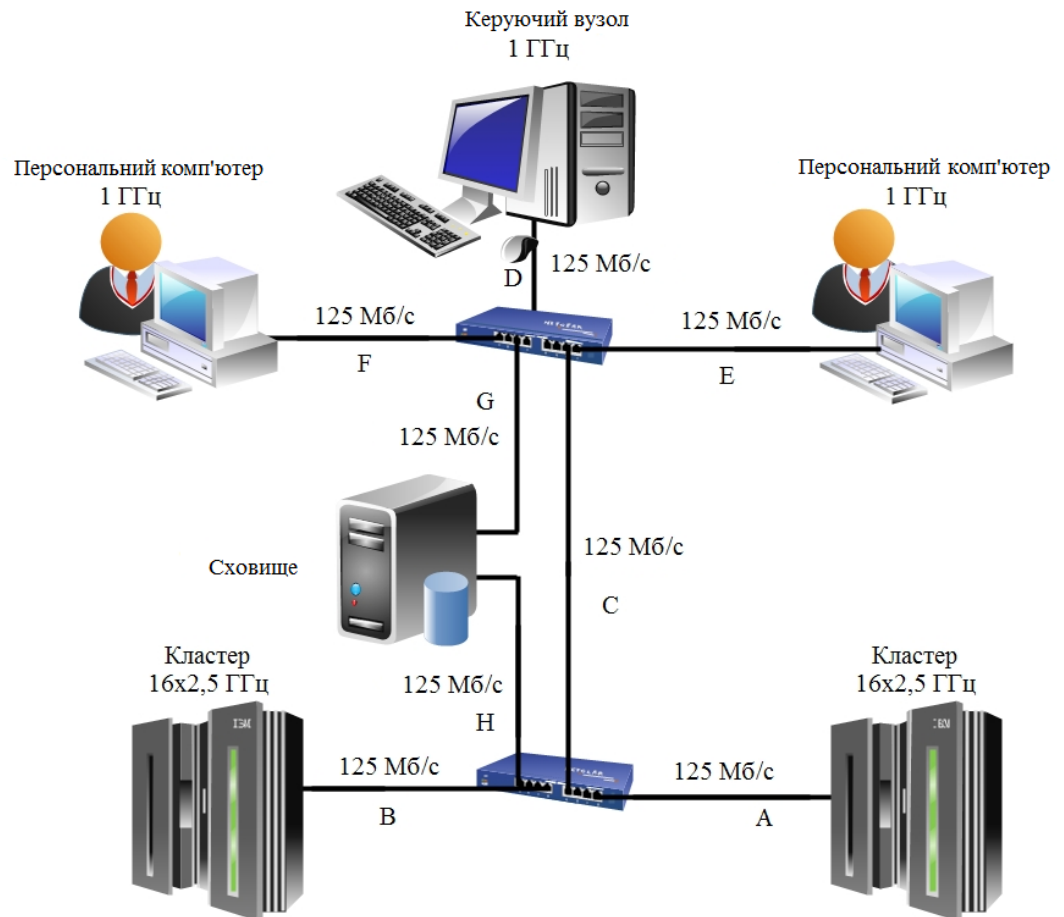


Рис. 4.4 Мережева топологія СОС зі сховищем

Відповідно, додатково кожену секунду необхідно передавати: $10^6 y_1$ на ділянці А, $10^6 y_2$ на ділянці В, $10^6(y_1 + y_2)$ на ділянці С, $10^6 x_1$ на ділянці Е, $10^6 x_2$ на ділянці F та $10^6(x_1 + x_2 + y_1 + y_2)$ на ділянці D.

Таким чином, загальне навантаження на мережу становить:

$(10^6 + 5 \cdot 10^8)u_1$ на ділянці А, $(10^6 + 5 \cdot 10^8)u_2$ на ділянці В, $10^6(u_1 + u_2)$ на ділянці С, $10^6(x_1 + x_2 + u_1 + u_2)$ на ділянці D, $(10^6 + 5 \cdot 10^8)x_1$ на ділянці Е, $(10^6 + 5 \cdot 10^8)u_2$ на ділянці F, $5 \cdot 10^8(x_1 + x_2)$ на ділянці G, $5 \cdot 10^8(u_1 + u_2)$ на ділянці Н.

При цьому на величини x_1 та x_2 накладаються обмеження: їх значення не можуть перевищувати кількість задач, які можна виконати на одному персональному комп'ютері за одну секунду $x_{1,2} \leq 3,70 \cdot 10^9 / 10^9 = 3,7$. Для кластерів таке обмеження має вигляд: $u_{1,2} \leq 7,27 \cdot 10^{10} / 10^9 = 72,7$.

Таким чином, поставлена задача звелася до наступної задачі лінійного програмування.

Необхідно максимізувати функцію $x_1 + x_2 + u_1 + u_2$, при наступних обмеженнях на x_1, x_2, u_1, u_2 :

$$\begin{aligned} (10^6 + 5 \cdot 10^8)u_1 &\leq sp(A), \\ (10^6 + 5 \cdot 10^8)u_2 &\leq sp(B), \\ 10^6(u_1 + u_2) &\leq sp(C), \\ 10^6(x_1 + x_2 + u_1 + u_2) &\leq sp(D), \\ (10^6 + 5 \cdot 10^8)x_1 &\leq sp(E), \\ (10^6 + 5 \cdot 10^8)x_2 &\leq sp(F), \\ 5 \cdot 10^8(x_1 + x_2) &\leq sp(G), \\ 5 \cdot 10^8(u_1 + u_2) &\leq sp(H), \\ x_1 &\leq 3,7, \\ x_2 &\leq 3,7, \\ u_1 &\leq 72,7, \\ u_2 &\leq 72,7, \\ x_1 \geq 0, x_2 \geq 0, u_1 \geq 0, u_2 \geq 0. \end{aligned}$$

Нехай пропускна здатність мережі на всіх ділянках залишає 1000 Мбіт/с=125 Мб/с. З рис. 4.4 маємо $x_1 + x_2 = u_1 + u_2 = 0,25$. Тоді, продуктивність СОС становить $x_1 + x_2 + u_1 + u_2 = 0,25$ задач/с. При цьому настільки низька продуктивність пояснюється великим обсягом даних, які необхідно пересилати по мережі для виконання одного завдання [187].

Експеримент 2.

Оцінка ефективності передачі декількох файлів даних меншого розміру в сервіс-орієнтованій системі

Розглянемо тепер випадок, коли для виконання задач потрібно завантажити зі сховища не один файл великого об'єму, а декілька файлів меншого розміру. Припустимо, що в результаті виконання задачі на обчислювальному вузлі генеруються файли, які необхідно завантажити в сховище. Таким чином, внесемо наступні корективи:

1. Розмір файлів, які передаються по мережі, становлять 1 Мб. При цьому розмір виконуваного коду (кількість необхідних для виконання машинних інструкцій) становить 1000 Мб.

2. Для розв'язання задач на обчислювальний вузол необхідно завантажити зі сховища 10 файлів розміром 25 Мб кожен (всього 250 Мб).

3. Після виконання задачі отриманий результат необхідно зберегти у сховище у вигляді одного файлу розміром 250 Мб.

4. Швидкість читання зі сховища при зчитуванні 1, 2, 3, ... файлів становить відповідно 450, ..., 200 (десятий член послідовності), ... Мб / с.

5. Швидкість запису в сховище, при записі 1, 2, 3, ... файлів становить відповідно 125, 110, 100, 95, ... Мб / с.

Потрібно. Визначити продуктивність сховища при зчитуванні з нього 10 файлів невеликого розміру.

Так як в цьому випадку мова йде про зчитування 10 файлів, то продуктивність жорсткого диска (або дисків) сховища становить 200Мб/с:

$$f_{HDD} = 2 \cdot 10^8 \text{ (байт/с)}, \text{ маємо } f_{read} = 2 \cdot 10^8 \text{ (байт/с)}.$$

Аналогічно знайдемо продуктивність сховища при завантаженні файлу великого розміру. У цьому випадку (тобто при записі одного файлу) продуктивність диска сховища становить 125 Мб / с або $f_{HDD} = 1,25 \cdot 10^8 \text{ (байт/с)}$. Тобто $f_{write} = 1,25 \cdot 10^8 \text{ (байт/с)}$.

Таким чином, продуктивність сховища при читанні є достатньою для

забезпечення регулярного зчитування (вивантаження в мережу) 200 Мб/с і запису (завантаження з мережі) тільки 125 Мб / с. В обох випадках це не перевищує пропускної здатності з'єднання сховища і мережі (2 з'єднання по 125 Мб / с кожне).

Нехай, в середньому, на персональні комп'ютери кожную секунду подається x_1 та x_2 задач. На кластери відповідно подається y_1 та y_2 задач за секунду. Тоді для забезпечення розв'язання цих задач кожную секунду необхідно завантажувати зі сховища $2,5 \cdot 10^8(x_1 + x_2)$ байт на маршрутизатор, показаний на рис. 4.4 зліва, і $2,5 \cdot 10^8(y_1 + y_2)$ на маршрутизатор праворуч. В результаті розв'язання задачі від маршрутизатора зліва надходить $2,5 \cdot 10^8(x_1 + x_2)$ байт, а від маршрутизатора праворуч — $2,5 \cdot 10^8(y_1 + y_2)$ байт.

Решта обмежень можна отримати аналогічно попередньому прикладу. При цьому необхідно врахувати наступне: сховище має встигати зчитувати і записувати інформацію із заданою швидкістю. Обчислимо цю величину. При розв'язанні однієї задачі сховище вивантажує файли розміром 250 Мб зі швидкістю 200 Мб / с і завантажує файл розміром 250 Мб зі швидкістю 120 Мб/с. На вивантаження файлів йде $T_{out} = 250/200 = 1,25$ секунд, на завантаження $T_{in} = 250/125 = 2$ секунд, що в сумі становить 3,25 сек. Для забезпечення виконання $x_1 + x_2 + y_1 + y_2$ задач, сховище буде зайнято протягом $3,25 \cdot (x_1 + x_2 + y_1 + y_2)$ с, що повинно бути менше 1 с.

Поставлена задача звелася до наступної задачі лінійного програмування: необхідно максимізувати функцію $x_1 + x_2 + y_1 + y_2$, при наступних обмеженнях на x_1, x_2, y_1, y_2 :

$$\begin{aligned}(10^6 + 5 \cdot 10^8)y_1 &\leq sp(A), \\ (10^6 + 5 \cdot 10^8)y_2 &\leq sp(B), \\ 10^6(y_1 + y_2) &\leq sp(C), \\ 10^6(x_1 + x_2 + y_1 + y_2) &\leq sp(D), \\ (10^6 + 5 \cdot 10^8)x_1 &\leq sp(E),\end{aligned}$$

$$(10^6 + 5 \cdot 10^8)x_2 \leq sp(F),$$

$$5 \cdot 10^8(x_1 + x_2) \leq sp(G),$$

$$5 \cdot 10^8(y_1 + y_2) \leq sp(H),$$

$$x_1 \leq 3,7,$$

$$x_2 \leq 3,7,$$

$$y_1 \leq 72,7,$$

$$y_2 \leq 72,7,$$

$$3,25(x_1 + x_2 + y_1 + y_2) \leq 1$$

$$x_1 \geq 0, x_2 \geq 0, y_1 \geq 0, y_2 \geq 0.$$

Пропускна здатність мережі становить 1000 Мбіт/с=125 Мб/с, швидкість з'єднання на всіх ділянках мережі становить $1,25 \cdot 10^8$ (байт/с). Тоді, враховуючи припущення, що жорсткий диск такого комп'ютера має розмір 60 Гб, а кожен багатопроцесорний обчислювальний вузол містить 16 процесорів з частотою 2,5 ГГц. Розмір загальної оперативної пам'яті становить 2 Гб і ємність загального жорсткого диска - 500 Гб, після чого отримаємо наступне рішення $x_1 + x_2 \leq 0,25$, $y_1 + y_2 \leq 0,25$. Однак внаслідок припущення що продуктивність пам'яті (тобто швидкість читання / запису) складає 10 Гб / с, продуктивність складає $x_1 + x_2 + y_1 + y_2 = \frac{4}{13}$ задач/с. Така невелика продуктивність пояснюється великими обсягами даних, які передаються по мережі для виконання завдань і малою продуктивністю сховища [187]. Тому, в розробленій сервіс-орієнтованій системі всі дані зберігаються на одному сховищі для забезпечення максимальної продуктивності розв'язання задачі картографування земного покриття (що було досліджено та експериментально підтверджено в пункті 4.2-4.3).

4.4 Візуалізація отриманих результатів картографування земного покриття

Наступним етапом є візуалізація отриманих результатів. Базовими вимогами до даної підсистеми є створення програми, яка б дозволила

користувачеві з будь-яким рівнем володіння комп'ютером отримувати доступ до даних, які йому необхідні. Також, доступ до даних, розміщених на порталі, повинен бути максимально оперативним, швидким і зручним.

Для забезпечення постійного доступу до даних пропонується створення Web-сервісу, який буде містити функції, необхідні для користувачів. Web-сервіс не вимагає установки на комп'ютер, не залежить від операційної системи і всі складні обчислення проводяться центральним сервером, що робить можливим доступ до даних практично з будь-якого комп'ютера в світі.

Для надання результатів роботи користувачам в сервісі реалізовані сучасні засоби публікації результатів обробки геопросторових даних в мережі Інтернет. В якості таких засобів вибрано безкоштовне програмне забезпечення з відкритим кодом QGIS-сервер для відображення побудованих карт земного покриття. Наприклад, карта класифікації Київської області на геопорталі (рис. 4.5).

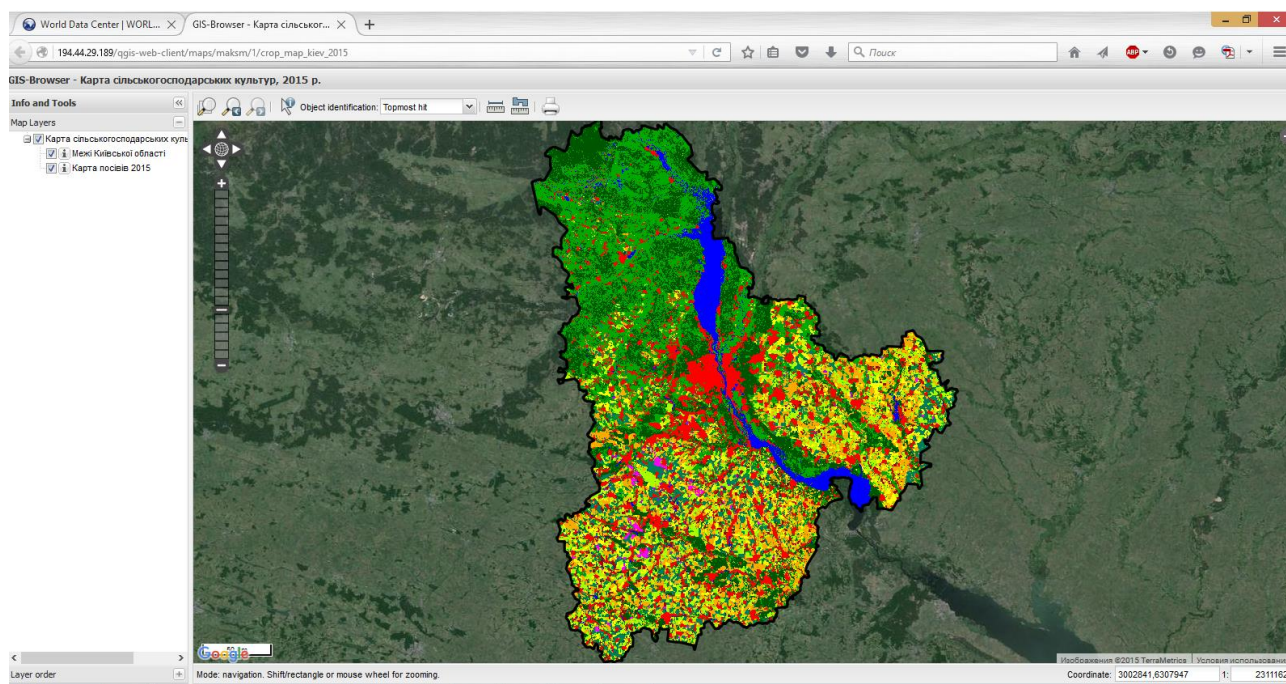


Рис. 4.5 Карта класифікації Київської області на геопорталі

Для створення геопортала було обрано MVC Framework, так як він безкоштовний, надійний, активно розвивається, безпечний і зручний. Framework пов'язує базу даних з Html-сторінками, які переглядає користувач.

Наявність в ньому великої кількості корисних готових рішень дозволяє прискорити процес розробки і зменшити обсяг необхідного коду, який підвищує його надійність.

Для відображення даних на стороні клієнта використаний QGIS Web-клієнт - це Java-клієнт для зручного відображення інформації, одержуваної з QGIS Mapserver.

Геопортал містить дві бази даних - створені на базі PostgreSQL і PostGIS DB, що дозволяє легко зберігати і швидко шукати необхідну інформацію за географічними даними. Пошук за географічними даними не є простим завданням для бази даних, адже індекс для швидкого знаходження необхідного об'єкта не є тривіальним. Для побудови цього індексу і застосовується PostGIS DB. Для всіх інших даних доступні всі функції PostgreSQL. Для роботи з даними використовуються наступні бібліотеки: PROJ, GEOS, GDAL. Вони забезпечують зручний інтерфейс для роботи з географічними даними.

4.5 Оцінка репутації ресурсів в сервіс-орієнтованій системі

При виборі сервіс-орієнтованої системи, важливою складовою є оцінка репутації ресурсів. Репутація - це припущення про очікувану надійність ресурсу за існуючою інформацією або спостереженнями за його поведінкою в минулому [16, 189]. Для цього пропонується інтегрувати репутацію в розроблену автоматизовану інформаційну технологію в сервіс-орієнтовану систему (схема на рис. 4.6). Розглянемо основні компоненти запропонованої схеми:

- **Користувачі.** Використовують сервіси. Може бути особою, яка одержує доступ до сервісів, наприклад, через портал, так і сторонніми системами, які одержують доступ до сервісів через програмні інтерфейси. Крім того, користувачі формують віртуальну організацію (ВО).
- **Ресурс.** Надає сервіси і формує ВО.
- **Віртуальна організація.** Тимчасове або постійне об'єднання географічно

розподілених окремих осіб, груп, підрозділів організацій або цілих організацій, які спільно використовують інформацію та ресурси для досягнення спільних цілей.

- **Сервер VOMS.** Система обліку VO.
- **Сервіс.** Функціональність, яка надається ресурсом і компонентами системи. У сервіс-орієнтованих системах будь-яка взаємодія між компонентами здійснюється за допомогою сервісів.
- **Каталог сервісів.** Містить метадані про сервіси системи.

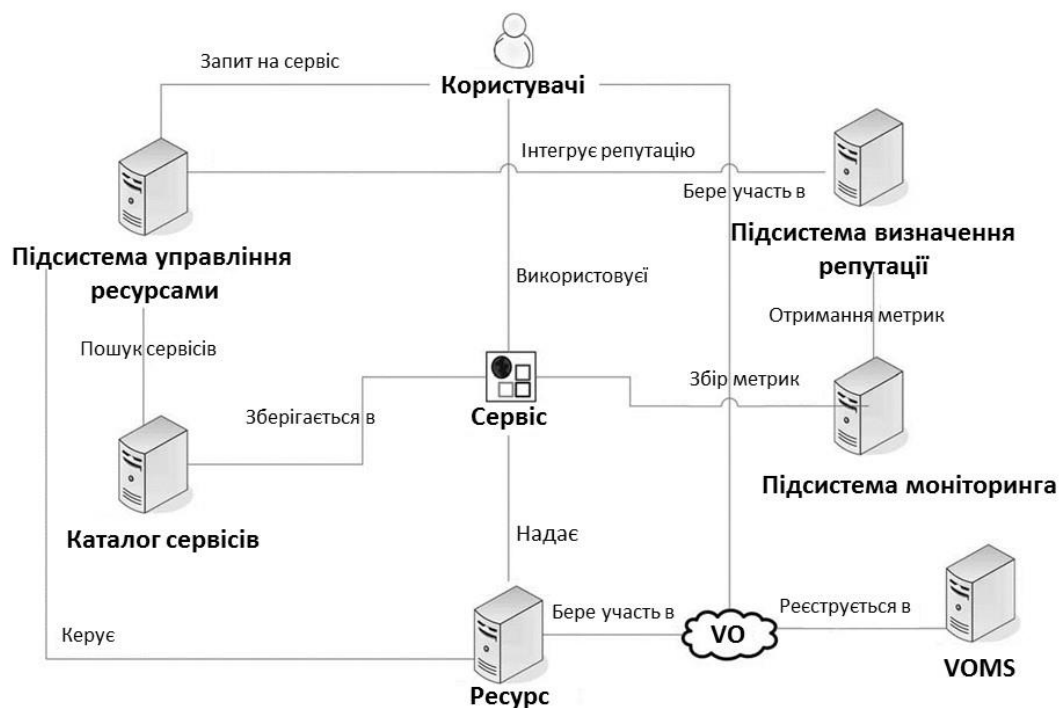


Рис. 4.6 Інтеграція репутації інформаційної технології в сервіс-орієнтовану систему

- **Підсистема керування ресурсами.** Одержує запити від користувачів, обробляє їх, здійснює пошук у каталозі сервісів, що задовольняють вимогам користувача, і вибирає необхідний ресурс для надання сервісу. Для вибору ресурсів використовується репутація ресурсу, яка отримується з підсистеми керування репутацією.
- **Підсистема визначення репутації.** Реалізується на основі інформаційної технології і обчислює, зберігає і оновлює репутацію ресурсів. Взаємодіє з підсистемою моніторингу для одержання значень метрик якості

обслуговування для обчислення функції корисності і репутації [16]. Взаємодіє з підсистемою керування ресурсами для інтеграції репутації в планувальник ресурсів.

- **Підсистема моніторингу.** Збирає і зберігає інформацію про взаємодію ресурсу і користувача.

Таким чином, використання репутації ресурсу для вибору сервісу дозволяє виключити або обмежити використання ненадійних постачальників сервісів.

4.6 Висновки до розділу

У четвертому розділі зосереджена увага на розробці автоматизованої інформаційної технології картографування земного покриття для великих територій у вигляді сервіс-орієнтованої системи. Проведені дослідження, що дозволили сформулювати прямі та зворотні задачі оптимізації СОС і її окремі компоненти. Проаналізовано результати експериментальних досліджень для оцінки ефективності СОС різної апаратної конфігурації і різних функціональних елементів. В результаті використання СФА для оцінки ефективності передачі одного і декількох файлів для однієї з архітектур на основі СОС найбільш вузьким місцем виявилось сховище. Тобто, для забезпечення ефективності роботи автоматизованої інформаційної технології геопросторів дані та програмний код необхідно зберігати та обробляти на одному сховищі або переходити до хмарних технологій, зокрема Google Earth Engine.

Таким чином, на основі розглянутого підходу можна виконувати аналіз СОС-архітектури і виявляти їх «вузькі» місця. Проведені дослідження дозволили підвищувати ефективність роботи автоматизованої інформаційної технології побудови карти земного покриття для великих територій. Для зручності візуалізації результатів обробки розроблено геопортал на основі безкоштовного програмного забезпечення з відкритим кодом QGIS-сервер.

РОЗДІЛ 5

ПРАКТИЧНІ ЗАСТОСУВАННЯ РОЗРОБЛЕНИХ МЕТОДІВ ЗЛИТТЯ ДАНИХ ТА АВТОМАТИЗОВАНОЇ ІНФОРМАЦІЙНОЇ ТЕХНОЛОГІЇ КАРТОГРАФУВАННЯ ЗЕМНОГО ПОКРИВУ

В даному розділі описані прикладні задачі, до яких застосовуються розроблені методи злиття даних з використанням автоматизованої інформаційної технології картографування земного покриття для різних областей території України. Практичне застосування розроблених методів дали можливість здійснювати оцінку площ кожного із класів земного покриття, отримувати об'єктивні оцінки збитків в наслідок посухи, можливість оцінити зміни земного покриття протягом кількох десятиліть та ін.

5.1 Практична реалізація запропонованого методу злиття даних для території України

Однією з задач, яка досліджена в даному пункті дисертаційного дослідження – це побудова карти земного покриття для всієї території України за три десятиліття (1990, 2000 та 2010 роки) з використанням розроблених у другому розділі методів злиття супутникових даних. Для розв'язання поставленої задачі використовувались супутникові знімки Landsat-4/5/7 (табл. 3.11) та вибірка, що була побудована по фотоінтерпретації (табл. 3.12). Вибірка містить 6 класів: штучні об'єкти, землі с/г призначення, ліси, необроблювані землі, відкритий ґрунт та вода. На рисунках (рис 5.1-5.3) представлені карти для 2010, 2000 та 1990 років. Ці карти дозволяють оцінити загальні тенденції різного призначення земного покриття в Україні. Наприклад, порівняння карт для 1990, 2000 і 2010 показали збільшення необроблюваних земель, зокрема в північній частині України. Основною причиною є те, що багато полів були закинуті через їх низьку врожайність, і культивування цієї

землі було не вигідно і не доцільно для фермерів.

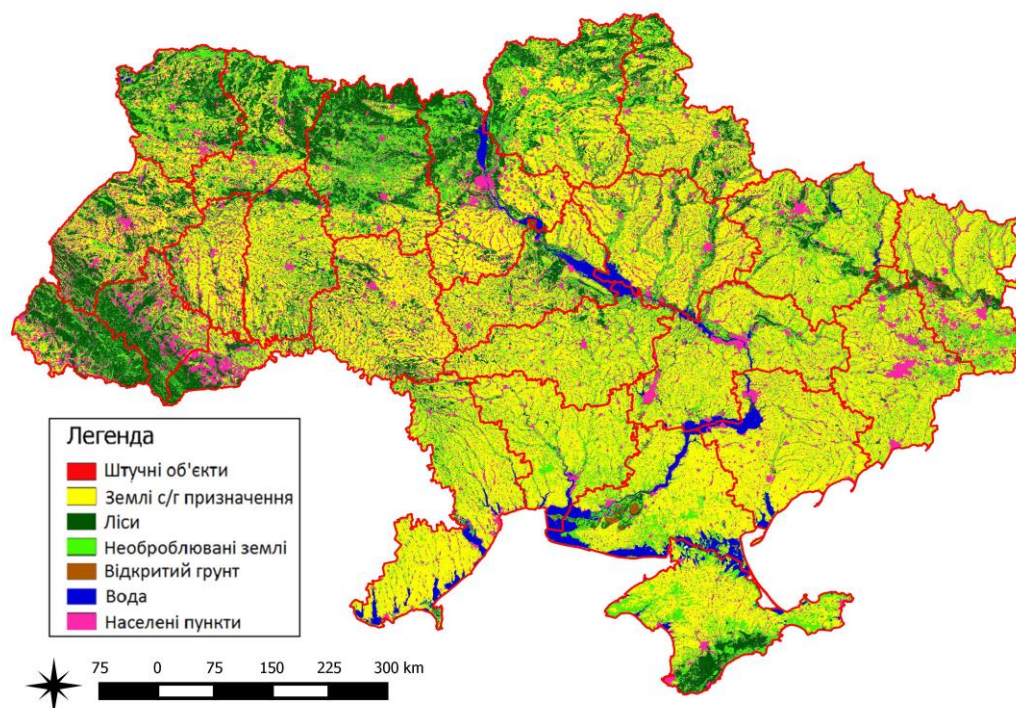


Рис. 5.1 Карта класифікації культур для всієї території України, використовуючи супутникові знімки Landsat 5/7 в 2010 році

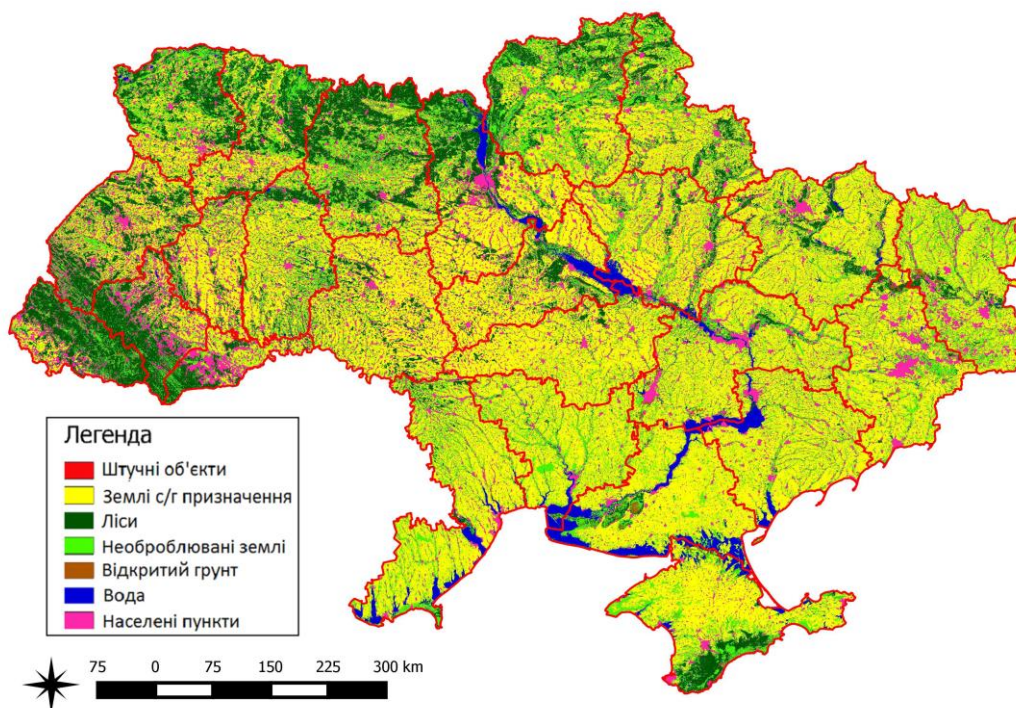


Рис. 5.2 Карта класифікації культур для всієї території України, використовуючи супутникові знімки Landsat 5/7 в 2000 році

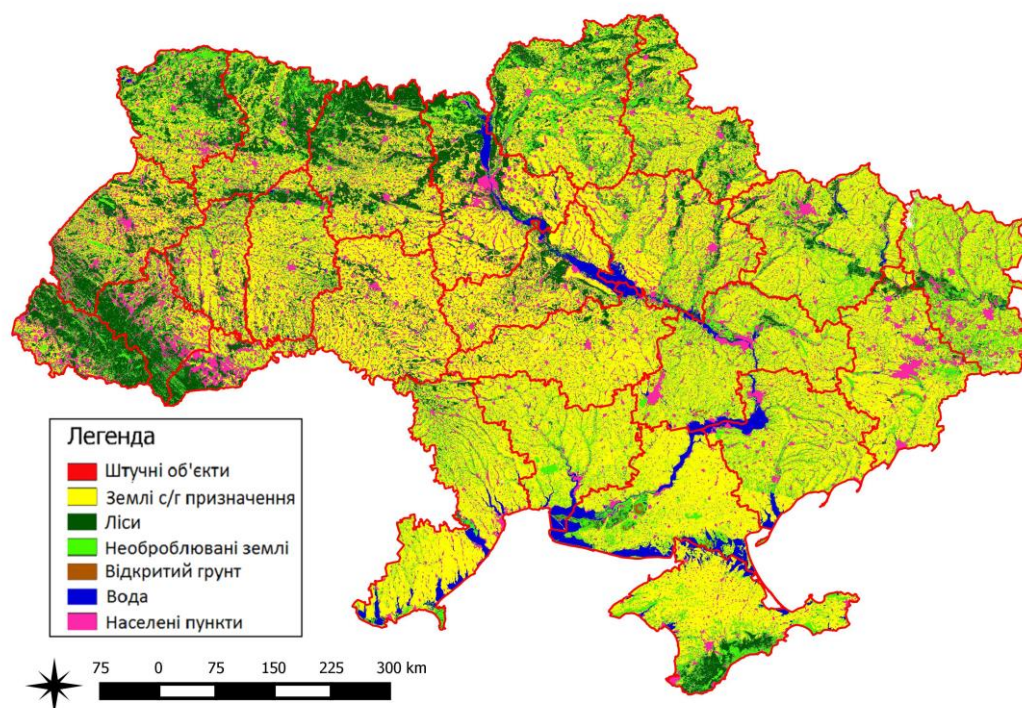


Рис. 5.3 Карта класифікації культур для всієї території України, використовуючи супутникові знімки Landsat 4/5 в 1990 році

Точність класифікації земної поверхні (табл. 3.13) напряму залежить не тільки від якості використаних даних, а й від їхньої кількості. У випадку, коли для досліджуваної території отримано відповідний часовий ряд зображень, є можливість враховувати динаміку змін земної поверхні, і таким чином підвищувати якість інтерпретації даних. Це дозволяє точніше ідентифікувати такі типи земної поверхні, як необроблювані землі, воду (виокремити постійні водойми від тимчасово затоплених територій), відділити озимі культури від ярих тощо. Середня точність класифікації земного покриття по всій території складає 93.6%.

5.2 Карти змін

5.2.1 Карти змін земного покриття для території України

В даному пункті практичною задачею є провести геопросторовий аналіз змін земного покриття на основі побудованих карт земного покриття для всієї

території України протягом трьох десятиліть (представлених в пункті 5.1). Побудовані карти проаналізовані на трьох рівнях (піксельному, в розрізі адміністративних районів та областей). Графічне представлення отриманих результатів наведено на рис. 5.4. На растровій (попіксельній) карті (рис. 5.4 зліва) показані основні переходи: ліловим кольором виділено перехід оброблюваних земель в необроблювані, зеленим – необроблюваних в оброблювані. На рис. 5.4 наводяться результати узагальнення по районам (посередині) і областям (справа). Червоним кольором показані адміністративні одиниці, для яких збільшилася площа необроблюваних земель, а зеленим кольором виділені ті з них, для яких вона зменшилася.

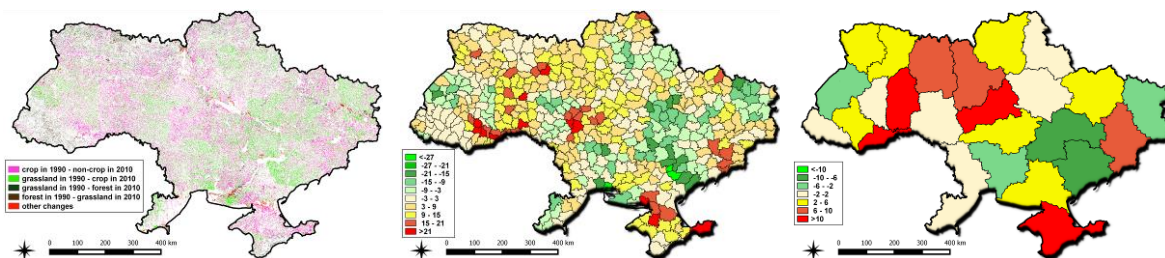


Рис. 5.4 Зміни земного покриття України з 1990 по 2010 р.

Порівнюючи отримані результати для областей України за період з 1990 по 2010 р. з даними офіційної статистики, можна побачити певні розбіжності. Порівняльний аналіз зміни площі оброблюваних земель для окремих областей за період 1990 – 2010 рр. за статистичними даними і на основі супутникових спостережень наводиться на рис. 5.5.

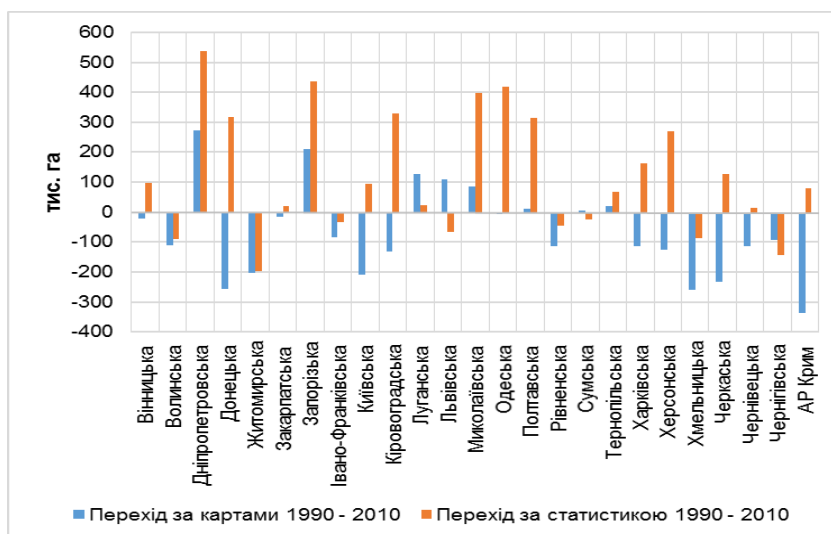


Рис. 5.5. Зміни площі оброблюваних земель в період 1990-2010 рр. за статистичними даними та по картах класифікації

Для більш детального аналізу аналогічно побудовані карти змін і за періоди 1990-2000 і 2000-2010 роки (рис.5.6).

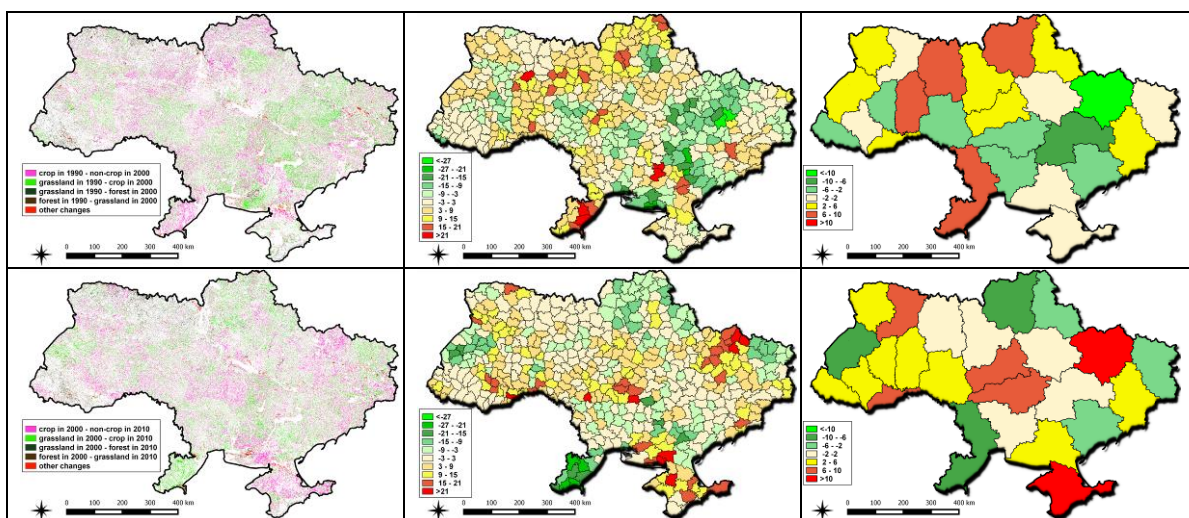


Рис. 5.6 Зміни земного покриття України з 1990 по 2000 (зверху) і з 2000 по 2010 рр. (внизу)

Отримані результати свідчать про те, що в північно-західних областях України помітною тенденцією є збільшення площ необроблюваних земель. Найбільші площі закинутих земель спостерігаються в Чернівецькій, Хмельницькій, Черкаській областях та в АР Крим. Дещо краща ситуація в Житомирській та Київській областях. І лише у Львівській, Миколаївській, Луганській, Запорізькій та Дніпропетровській областях спостерігається збільшення оброблюваних сільськогосподарських земель.

5.2.2 Карти змін земного покриття Донецької та Луганської областей

В даному пункті, використовуючи отримані результати (в пункті 5.2.1) та добудувавши карту земного покриття для 2015 року, оцінимо зміни земного покриття і землекористування в Донецькій та Луганській областях, особливо на окупованих територіях. Для отримання результатів були використані дані супутника Proba-V (для 2015 р.) та Landsat-5/7 (для 2000 та 2010 рр.). На рис. 5.7-5.8 наведені карти класифікації для Донецької та Луганської областей,

отримані на основі запропонованого у другому розділі підходу. Чорним кольором на рисунках показано межі окупованої території.

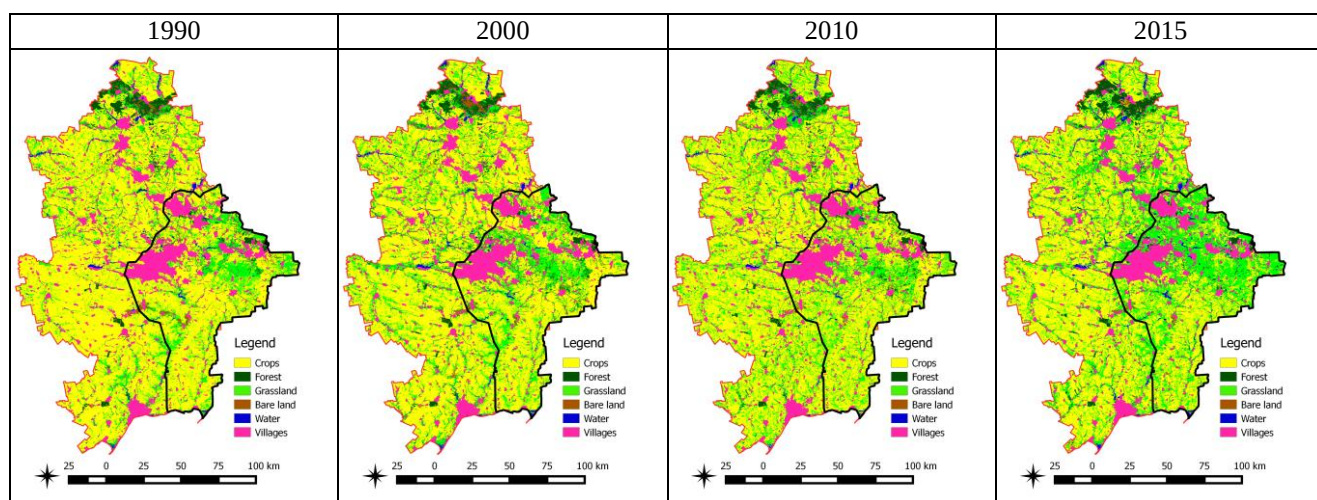


Рис. 5.7 Карти земного покриття для Донецької області

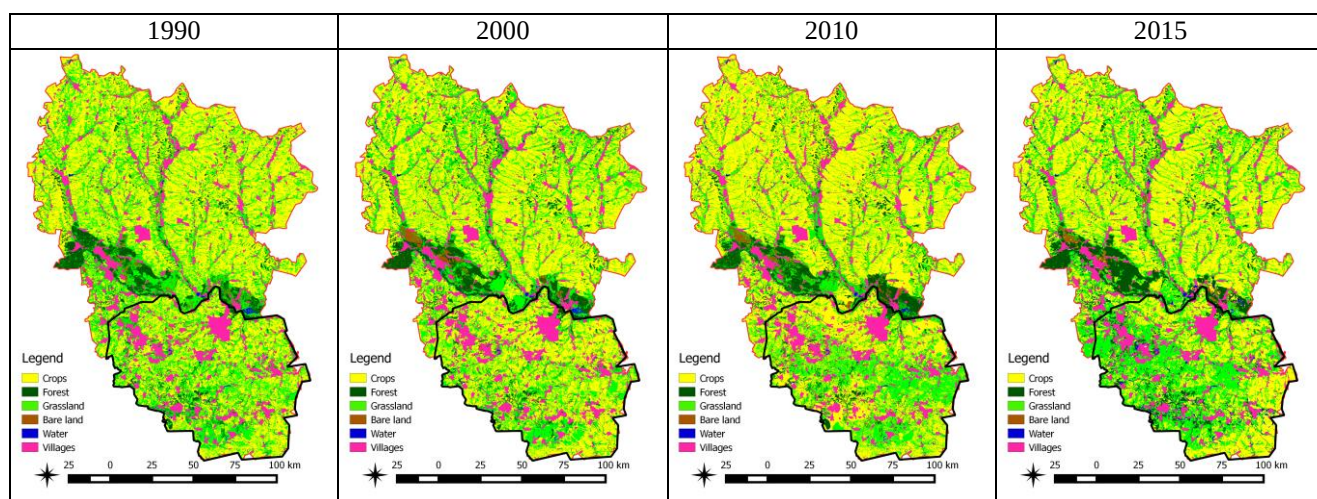


Рис. 5.8 Карти земного покриття для Луганської області

Як видно з рис. 5.7-5.8, максимальна площа оброблюваних земель в Донецькій області була у 1990 році, а в Луганській у 2000 році. В 2015 році площі оброблюваних земель істотно зменшилися. Як видно з табл. 5.1 і 5.2, в Донецькій області площа оброблюваних земель зменшилася більше ніж на 20% відносно 1990 року, причому значні негативні зміни відбулися переважно на окупованій території - 35%.

Для Луганської області також спостерігається істотне зменшення оброблюваних площ, близько 2% для всієї області і більше 15% на окупованих територіях.

Таблиця 5.1 –

**Площа оброблюваних і необроблюваних земель на окупованих територіях
Донецької та Луганської областей (1000 га)**

Клас \ Рік	Донецька область				Луганська область			
	2015	2010	2000	1990	2015	2010	2000	1990
Оброблювані землі	298,3	428,7	420,8	461,5	278,0	345,2	393,7	336,4
Необроблювані землі	347,3	208,8	211,6	179,3	336,2	285,1	245,4	299,2

Таблиця 5.2 –

**Площа оброблюваних і необроблюваних земель для
Донецької та Луганської областей (1000 га)**

Клас \ Рік	Донецька область				Луганська область			
	2015	2010	2000	1990	2015	2010	2000	1990
Оброблювані землі	1308,9	1522,2	1563,2	1695,5	1223,9	1396,4	1330,0	1257,4
Необроблювані землі	814,9	605,2	551,5	435,9	849,9	681,2	778,6	864,4

На рис. 5.9 представлений стан земного покриву на окупованих територіях.

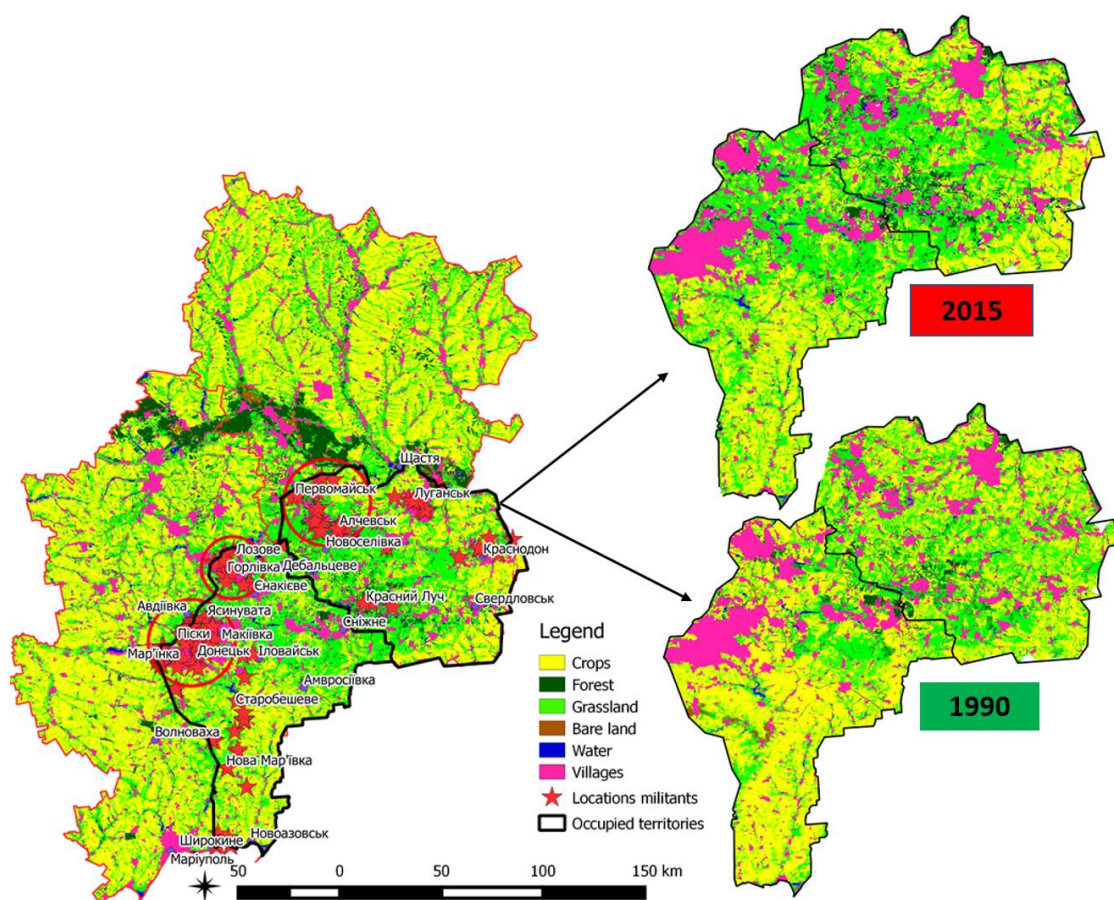


Рис. 5.9 Околиці найбільш постраждалих міст Донецької і Луганської областей

Червоними зірочками виділені населені пункти, в яких були зосереджені основні активні бойові дії. З рис. 5.9 видно, що в околицях цих населених пунктів спостерігається істотне зменшення площ оброблюваних сільськогосподарських земель.

Аналізуючи отримані площі, можна зробити висновок, що кількість оброблюваних земель для східних областей (Донецької та Луганської) в 2015 р. стало найменшим за період 1990-2015 рр. У той же час найбільшим є зменшення посівних площ на окупованих територіях, що складає 35% і 20% відносно 1990 року, відповідно (рис. 5.10).

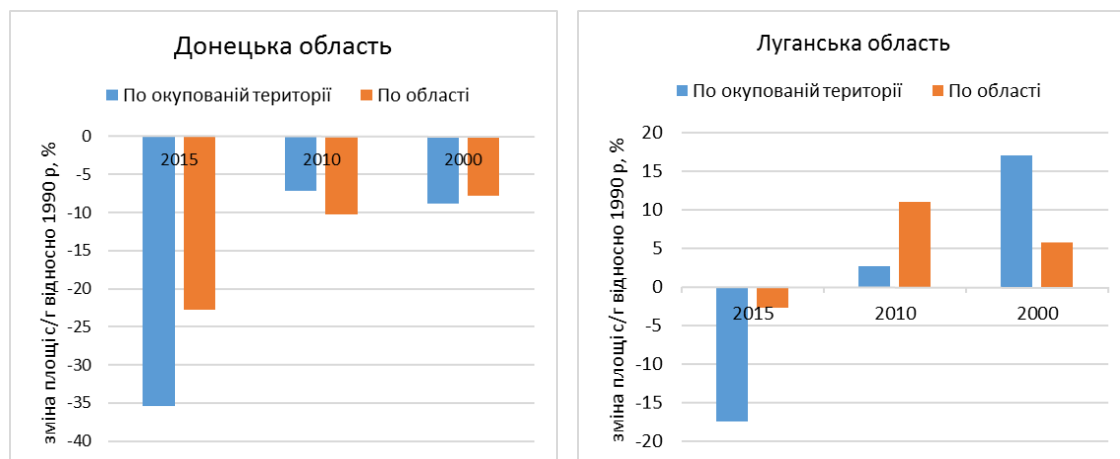


Рис. 5.10 Зміна площ оброблюваних земель в Донецькій і Луганській областях відносно 1990 року

В результаті проведеного дослідження можна зробити висновок, що внаслідок бойових дій, які відбуваються в східних областях України, частина полів Донецької та Луганської областей залишилася необробленою.

5.3 Порушення сівозміни

Для вирішення задачі порушення сівозміни на основі розроблених методів (описаних в другому розділі та експериментально перевірених в третьому) побудовано карти земного покриття для території Київської області. Моніторинг порушень сівозміни був проведений протягом 2013-2015 років (рис. 5.11). Представлені карти, побудовані на основі безкоштовних

супутникових даних із високим просторовим розрізненням, зокрема для 2015 р. використані Landsat-8 та Sentinel-1 (табл. 3.7). В якості навчальних та тестових даних використані дані наземних досліджень (табл. 3.8).

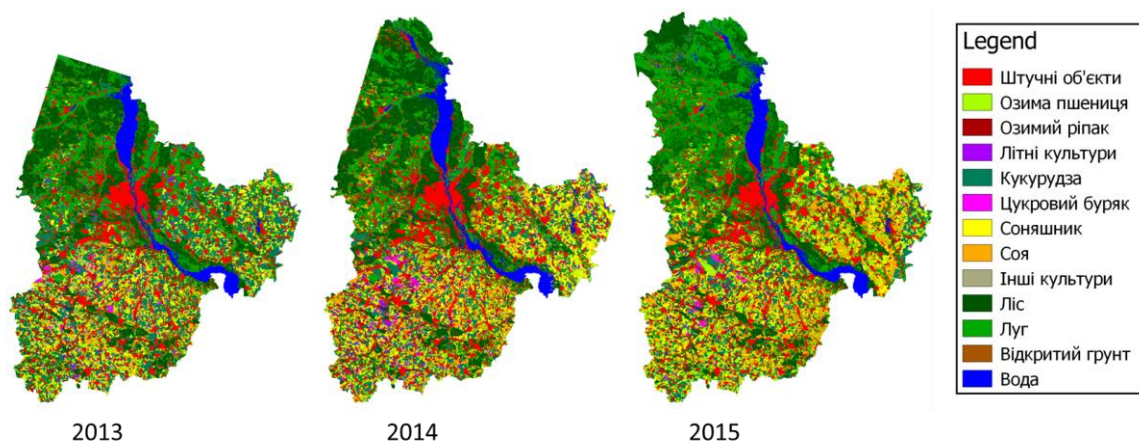


Рис. 5.11 Карти класифікації Київської області

На основі аналізу класифікації по супутниковим даним за 2013-2015 роки була створена тематична карта порушень сівозміни (рис. 5.12).

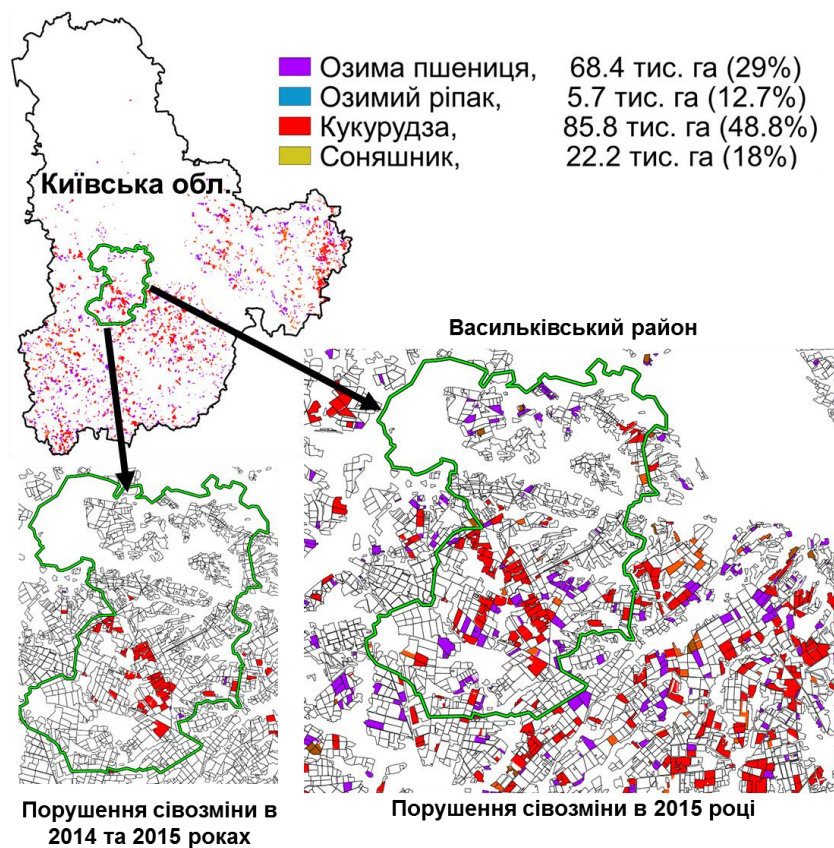


Рис. 5.12 Тематична карта порушень сівозміни

На карті фіолетовим кольором зображені порушення сівозміни озимої пшениці в 2015 році. Червоним кольором відображені території, на яких порушувалась сівозміна кукурудзи в 2014 та 2015 роках. Побудована карта дає можливість ідентифікувати поля на яких виснажується земля в результаті порушення сівозміни.

5.4 Аналіз збитків від посухи на основі злиття даних

Однією із задач, що розв'язуються за допомогою супутникових даних є моніторинг посухи. Будучи одним з найбільших стихійних лих, що впливають на життєдіяльність людства, посухи завдають колосальної шкоди багатьом сферам економіки, екології і т.д. [183, 190, 191]. Традиційно моніторинг посухи здійснюється на основі даних про опади, але отримання цих даних найчастіше виявляється проблематичним [192, 193]. Використання різноманітних даних дозволяє підвищити оперативність отримання інформації щодо стану сільськогосподарських культур, а також забезпечити максимальний рівень об'єктивності даних для своєчасного виявлення, моніторингу розвитку, оцінки збитків та зменшення наслідків посухи [194, 195]. Існують методи цифрової обробки космічних зображень, що дозволяють виділити території, постраждалі від посухи, і оцінити пов'язані з нею збитки [196, 197]. Більшість таких методів виявлення посухи і оцінки збитку з використанням супутникових даних ґрунтуються на порівнянні зображень за різні вегетаційні періоди, але такий підхід не досконалий, оскільки стан рослинності може погіршуватися з причин, не пов'язаних з посухою.

Для оцінки збитків від посухи, даних одного супутника може виявитися недостатньо, тому що частина даних може бути захмарена або можуть бути спотворення, викликані іншими факторами. Таким чином, необхідно аналізувати інформацію, що надходить з різних джерел з різним часовим і просторовим розрізненням. Основним джерелом об'єктивної інформації про стан посівів на великій площі є супутникові дані [198, 199].

Тому, пропонується використовувати підхід для оцінки збитків від посухи, що базується на технології злиття супутникових даних різного просторового розрізнення, векторних даних, що містять межі полігонів та даних наземних досліджень (детально описано в пункті 2.4.1).

5.4.1 Використані дані

Для зменшення витрат при розв'язанні задачі оцінки стану посівів пропонується використовувати безкоштовну інформацію супутників MODIS, Landsat та EO-1. Для оцінки площ, постраждалих від посухи, необхідні також векторні геопросторові дані з межами полігонів, які можна отримати або з відкритих джерел, або шляхом оцифровки по супутниковим знімкам середнього та високого просторового розрізнення [135, 200, 201]. Для формування навчальної вибірки при розв'язанні задачі класифікації посівів і їх пошкоджень необхідна наземна інформація про стан посівів на конкретних полях, для отримання якої необхідно залучати експертів (агрономів).

Оцінку збитків, заподіяних посухою, пропонується виконувати у два етапи. Спочатку по супутниковій інформації різного розрізнення з урахуванням векторних меж полігонів та наземних досліджень необхідно побудувати цифрові карти пошкоджень. Після цього можна перейти до оцінки площ пошкоджених посівів і економічних збитків.

5.4.2 Визначення пошкоджень сільськогосподарських культур на основі злиття даних

Для класифікації пошкоджень пропонується підхід, схема якого представлена на рис. 5.13 [125, 133]. Першим етапом обробки супутникових зображень для задачі аналізу збитків від посухи є кластеризація, яку пропонується здійснити за допомогою методу ISODATA (Iterative Self-Organizing Data Analysis Technique Algorithm). Наступним етапом обробки є

класифікація. Для її реалізації використовуються результати кластеризації зображення з подальшим віднесенням кожного кластера до деякого класу на основі наземної експертної інформації (навчальної вибірки, сформованої в результаті наземних досліджень) [202, 203]. Класифікація кожного знімка здійснюється за класом культур та пошкоджень.

З урахуванням специфіки задачі пропонується розділяти посіви на чотири класи: озимі, ярі, літні культури, необроблювані поля. Ступені пошкоджень пропонується розділити на три класи: неушкоджені посіви (до цього класу віднесемо ділянки, на яких постраждало менше 30% посівів), пошкоджені (знищено 30-70% посівів) і загиблі (втрати культур, що зійшли становлять понад 70%).

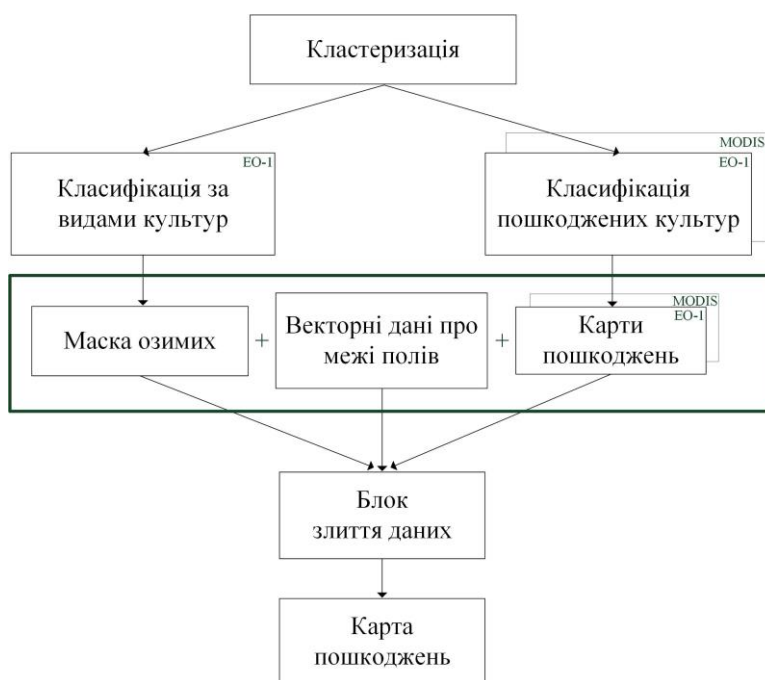


Рис. 5.13 Схема підходу до побудови карти пошкоджень

Для підвищення точності класифікації пошкоджень пропонується використовувати ансамбль моделей, призначених для обробки даних різних супутників (що розроблений в другому розділі). У блоці злиття даних на рис. 5.13 виконується геопросторовий аналіз інформації, отриманої в результаті роботи кожного конкретного класифікатора.

5.4.3 Оцінка збитків по картах класифікації

Економічний збиток, заподіяний посухою на деякій території, можна представити таким чином:

$$L = Y \cdot P \cdot S_{dm}, \quad (5.1)$$

де L – фінансовий збиток; Y – врожайність сільськогосподарської культури на даній території (отримується за даними статистики і вимірюється в тонн / га); P – ціна даної сільськогосподарської культури (наприклад, оцінка знизу за минулий рік, вимірювана в грн. / тонну), S_{dm} — площа пошкоджених (загиблих) посівів сільськогосподарської культур з урахуванням ступеня пошкодження.

Для оцінки площ S_{dm} загибелі посівів сільськогосподарських культур в (5.1) на основі супутникових даних можна скористатися наступною формулою:

$$S_{dm} = \sum_{i=1}^N \alpha_i \cdot S_i, \quad (5.2)$$

де N – число градацій пошкоджень; α_i , $i = \overline{1, N}$ – коефіцієнт пошкодження; S_i , $i = \overline{1, N}$ – сума площ пошкоджених посівів для кожного α_i .

5.4.4 Практична реалізація та експериментальна перевірка запропонованого підходу оцінки збитків від посухи

Експериментальна перевірка запропонованого методу визначення збитків виконувалася для Красногвардійського району АР Крим (в 2013 р.), на частку якого припадають найвищі показники за обсягом сільськогосподарського виробництва на півострові.

За допомогою Регіонального центру підтримки програми UN-SPIDER оперативно була організована космічна зйомка зазначеного району та отримані супутникові знімки EO-1 з просторовим розрізненням 30 м.

Визначення ступеня пошкодження полягає у візуальній оцінці кількості пошкоджених органів рослин [204]. Залежність між NDVI і ступенем

пошкоджень рослин можна апроксимувати регресійною залежністю, яка отримана з використанням даних наземних досліджень.

$$Damage = -0.0416 \cdot NDVI + 237.26, \quad (5.3)$$

Достовірність моделі (5.3) підтверджується коефіцієнтом детермінації $R^2 = 0.6687$.

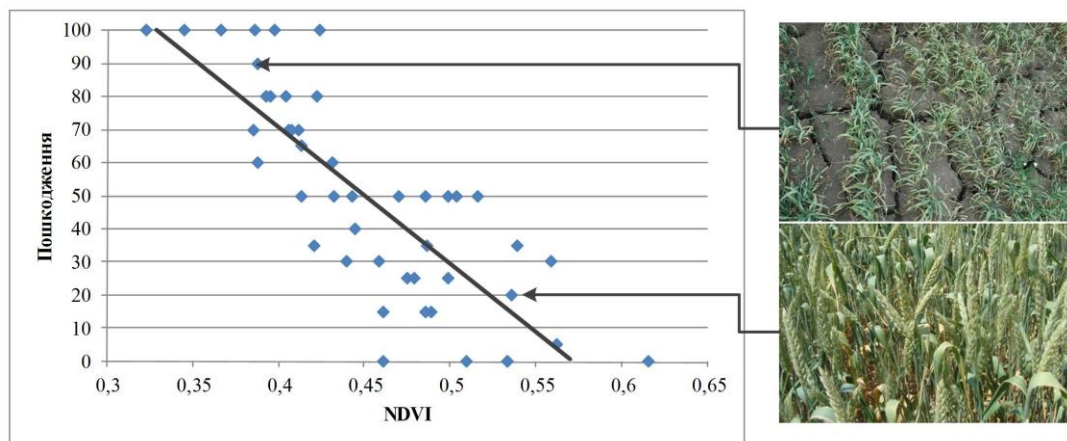


Рис. 5.14 Залежність значення NDVI від пошкоджень рослин

Наявність викидів на рис. 5.14 можна пояснити суб'єктивністю думки експерта, що проводив оцінку стану посівів. Для цього використано три можливі стани (неушкоджені, постраждалі, загинбі). Проте, такий підхід не дозволяє коректно класифікувати проміжні стани (переходи з одного стану в інший) посівів, а також ускладнює точну кількісну оцінку. Оскільки лінійна модель, побудована з використанням вищеописаних суб'єктивних експертних оцінок, погано описує наявні дані, прийнято рішення використовувати нелінійну модель, а саме застосувати класифікатор.

Оскільки карти пошкоджень в даному випадку будувалися за даними супутників EO-1 та Aqua (MODIS), формула (2.31) має вигляд:

$$K_l = W_{d^{EO}} \cdot K_l^{EO} + W_{d^{MOD}} \cdot K_l^{MOD}.$$

Застосування запропонованої технології злиття даних, описаної формулами (2.29-2.33), дозволило підвищити точність класифікації до 95% за рахунок залучення даних з інших джерел для покриття частин знімка, де відсутні дані (внаслідок погодних умов). Окремі класифікатори виступають в

такій схемі як експерти, що забезпечують перехід від аналізу окремих ознак до оцінки ситуації в цілому [205].

Загальна карта пошкоджень представлена на рис. 5.15.

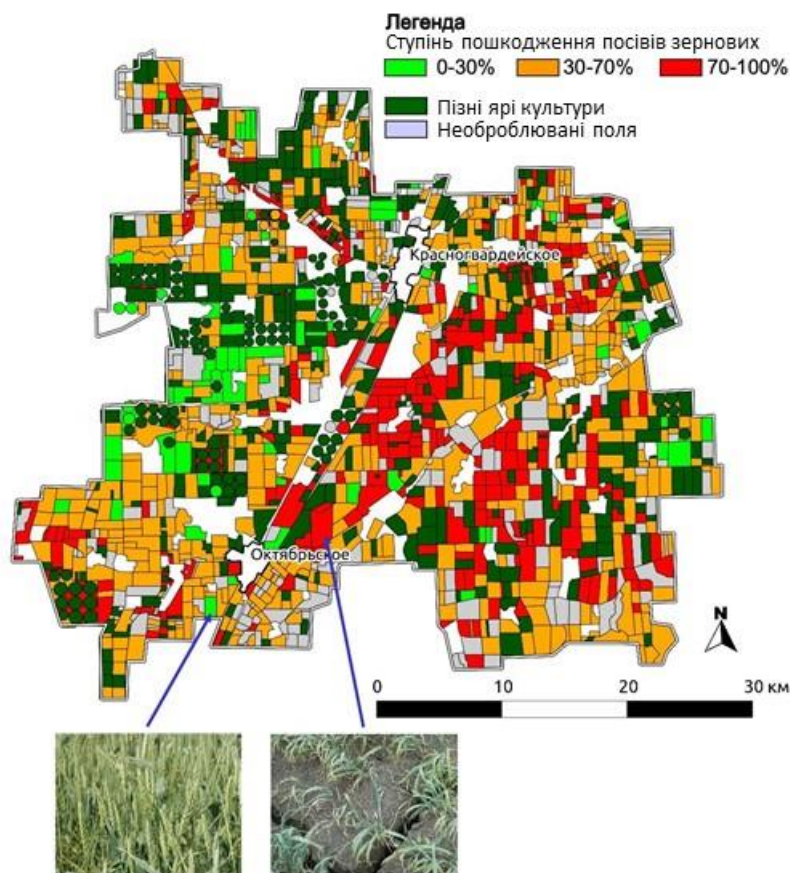


Рис. 5.15 Тематична карта стану посівів озимих та ярих зернових культур в АР Крим Красногвардійського району, 27 травня 2013

На основі отриманої геопросторової інформації про пошкодження посівів можна обчислити оцінку економічного збитку від посухи для Красногвардійського району за формулами (5.1-5.2). А саме, обчислимо площі посівів, віднесених до кожного з трьох розглянутих класів пошкоджень (рис. 5.15). Для співвідношення (5.2) отримаємо:

$S_1 = 8695$ га (площа посівів, пошкоджених на 0-30%);

$S_2 = 57250$ га (площа посівів, пошкоджених на 31-70%);

$S_3 = 25690$ га (площа посівів, пошкоджених на 71-100%).

Введемо коефіцієнти пошкодження для кожного з виділених класів.

Будемо вважати, що в (5.2) $\alpha_1 = 0$, $\alpha_2 = 0.5$ (посіви пошкоджені наполовину), $\alpha_3 = 1$ (урожай знищений повністю). Тоді, згідно (5.2), площа пошкоджених посівів (з урахуванням ступеня пошкоджень) становить:

$$S_{dm} = 8695 \cdot 0 + 57250 \cdot 0.5 + 25690 \cdot 1.0 = 54315.0 \text{ га}.$$

Враховуючи мінімальну ціну озимої пшениці та її середню врожайність для даного району, за формулою (5.2) можна обчислити оцінку економічного збитку по Красногвардійському району. Так, мінімальна ціна озимої пшениці в 2012 році становила 1700 грн. / Тонну [206], а згідно з даними сайту Головного управління статистики в Автономній республіці Крим врожайність озимої пшениці в Красногвардійському районі зафіксована на рівні 1,5 тонн / га [207]. Тоді відповідно до (5.1) отримаємо

$$L = 1.5 \text{ тонн / га} \cdot 1700 \text{ грн / тонну} \cdot 54315.0 \text{ га} = 138.5 \text{ млн. грн.}$$

Таким чином, в результаті несприятливих погодних умов економічний збиток по Красногвардійському району від загибелі зернових посівів у 2013 році складає близько 138.5 млн. грн., що узгоджується з оцінками інших експертів.

5.5 Висновки до розділу

В п'ятому розділі дисертаційного дослідження практично реалізовані розроблені методи злиття даних для побудови ретроспективних карт земного покриву для території України за супутниковими даними з високим просторовим розрізненням. На відміну від інших підходів [84, 85], запропонована методологія передбачає використання нейромережевої класифікації часових рядів супутникових даних, на яких попередньо відновлена інформація в захмарених областях, з подальшим об'єднанням отриманих карт для всієї території України. Використання інтелектуальних методів обробки часових рядів даних Landsat дозволило забезпечити загальну точність класифікації даних за 1990, 2000 та 2010 рр. на рівні 95% (на незалежній вибірці) і значно покращити якість картографування, порівняно з іншими

картами земного покриття України з просторовим розрізненням 30 м (пункт 5.1).

Також продемонстровано можливість побудови карти змін протягом кількох десятиліть (пункт 5.2). На основі запропонованих методів побудовані карти порушення сівозмін для території Київської області за 2013-2015 роки (пункт 5.3).

Застосування супутникової інформації та автоматизованих геоінформаційних технологій злиття даних дозволяє ефективно вирішувати оперативні завдання агромоніторингу для великих територій і отримувати об'єктивні оцінки збитків для підтримки прийняття управлінських рішень [183]. У розділі запропоновані методи оцінювання збитків від посухи на основі використання різномірної супутникової інформації. Визначення пошкоджених внаслідок посухи посівів виконується на основі геопросторового аналізу і злиття даних. Джерелами даних про пошкодження служать супутникові знімки з різних космічних апаратів, векторні дані із межами полігонів, отримані на наземних дослідженнях. Запропонований метод апробований на одному з районів Автономної Республіки Крим, який в 2013 році постраждав в результаті тривалої посухи. Аналіз тематичної карти показав, що в результаті несприятливих погодних умов (тривалої посухи) в Красногвардійському районі постраждали 54315 га посівів зернових (озимої пшениці та ярого ячменю), при цьому економічний збиток від загибелі посівів зернових по Красногвардійському району внаслідок несприятливих погодних умов склав близько 138.5 млн. грн. Надалі представлені методи і результати планується використовувати для створення системи підтримки прийняття рішень на регіональному рівні.

ВИСНОВКИ

У дисертаційному дослідженні розв’язана задача підвищення точності класифікації і завчасності оцінки площ класів земного покриття та забезпечення високої точності картографування земної поверхні на великих територіях шляхом розроблення методів злиття різнорідних геопросторових даних великих обсягів та створення автоматизованої технології картографування. Удосконалений метод класифікації земного покриття за супутниковими даними великих обсягів на основі використання ансамблю нейронних мереж та розроблені методи злиття гетерогенних супутникових даних на рівнях пікселів та прийняття рішень для побудови карти реалізовані у створеній автоматизованій інформаційній технології картографування для всієї території України в межах сервіс-орієнтованого підходу.

За результатами дисертаційної роботи можна зробити наступні висновки:

1. Проведений аналіз стану проблеми картографування земного покриття на великих територіях показав, що на сьогодні переважна більшість існуючих карт побудована на основі супутникових даних низького просторового розрізнення. Карти з високим просторовим розрізненням створено для невеликих територій, а існуюча єдина глобальна карта GlobeLand30-2010 характеризується низькою точністю і не враховує територіальних відмінностей, пов’язаних з кліматом та різнорідністю класів земного покриття. Тому актуальною є задача побудови карти земного покриття для великих територій, яку автор пропонує розв’язувати шляхом злиття різнорідних даних.

2. Розроблено метод злиття геопросторових супутникових даних на рівні пікселів, у якому розширюється простір ознак для картографування земного покриття шляхом інтеграції різночасових даних різної природи. Розроблений метод злиття оптичних та радарних геопросторових даних дозволив збільшити точність класифікації на 7,5% у порівнянні з випадком, коли використовуються лише оптичні дані.

3. Отримав подальший розвиток метод ансамблевої класифікації різнорідних геопросторових даних шляхом використання ансамблю нейронних мереж. Застосування ансамблю дозволяє підвищити точність класифікації земного покриття в порівнянні з окремою нейронною мережею більше ніж на 15% для ярих культур та на 6,5% для необроблюваних земель за рахунок об'єднання апостеріорних ймовірностей від окремих класифікаторів та обчислення середнього значення ймовірності належності вхідного образу до заданого класу.

4. Вперше розроблено методи злиття даних на рівні прийняття рішень для побудови карти земного покриття для великих територій. Використання розроблених методів злиття даних дозволило підвищити на 8% загальну точність класифікації земного покриття для території України порівняно з існуючою картою GlobeLand30-2010.

5. Розроблені методи злиття геопросторових даних реалізовані у формі сервіс-орієнтованої автоматизованої інформаційної технології. Проведені дослідження щодо оцінки ефективності сервіс-орієнтованої системи різної апаратної конфігурації і різних функціональних елементів. В результаті дослідження виявлено, що для забезпечення ефективності роботи автоматизованої інформаційної технології геопросторові дані та програмний код необхідно зберігати та обробляти на одному сховищі.

6. Використовуючи запропоновану автоматизовану технологію розв'язано задачу картографування земного покриття для всієї території України за 1990, 2000 і 2010 рік з 30-метровим просторовим розрізненням. Загальна точність класифікації для території України за областями становить не менше 88% і в середньому 97% при валідації на незалежній тестовій вибірці. Також вирішено проблеми оцінки збитків від посухи та порушення сівозміни. Побудовано карти земного покриття для території Київської області протягом 2013-2015 років. Для покращення достовірності карт земного покриття запропоновано метод злиття растрової та векторної інформації з використанням меж полігонів, що дає можливість покращити точність на 5,6% відносно

початкової карти класифікації.

Отримані в ході дисертаційного дослідження карти земного покриву включено до складу глобального геопросторового шару оброблюваних земель.

References

1. Kogan F. Winter wheat yield forecasting in Ukraine based on Earth observation, meteorological data and biophysical models / F. Kogan, N. Kussul, T. Adamenk, S. Skakun, O. Kravchenko, O. Kryvobok, A. Shelestov, A. Kolotii, O. Kussul, A. Lavrenyuk // *International Journal of Applied Earth Observation and Geoinformation*. — 2013. — Vol. 23, No 23. — P. 192–203. DOI:10.1016/j.jag.2013.01.002.
2. Куcсуль Н. М. Геопросторовий аналіз ризиків стихійних лих / Н. М. Куcсуль, С. В. Скакун, А. Ю. Шелестов // К.: «Наукова думка». — 2014. — 258 с.
3. Kussul N. Crop clasification in Ukraine using satellite optical and SAR images / N. Kussul, S. Skakun, A. Shelestov, O. Kravchenko, O. Kussul // *Models&Analyses*. — 2013. — Vol. 2, No. 2. — P. 118–128.
4. Gallego J. Efficiency assessment of using satellite data for crop area estimation in Ukraine / J. Gallego, N. Kussul, S. Skakun, O. Kravchenko, A. Shelestov, O. Kussul // *International Journal of Applied Earth Observation and Geoinformation*. — 2014. — Vol. 29. — P. 22–30.
5. Jixian Zhang. Multi-source remote sensing data fusion: status and trends / Jixian Zhang // *International Journal of Image and Data Fusion*. – 2010. – Vol. 1, No. 1.– P. 5–24.
6. Яйлимов Б. Я. Метод класифікації на основі слияння даних для аналізу ушкоди від засухи / Б. Я. Яйлимов // *Індуктивне моделювання складних систем*. — 2014. — Вип. 6. — С. 167–176.
7. Шелестов А. Ю. Інформаційна технологія оцінки ушкоди від засухи на основі слияння даних / А. Ю. Шелестов, Б. Я. Яйлимов, А. І Петухова // *Наукові праці Донецького національного технічного університету Серія «Інформатика, кібернетика та обчислювальна техніка»*. – 2013. – Т. 17, № 1. – С. 125–132.
8. Kussul N. Regional scale crop mapping using multi-temporal satellite imagery / N. Kussul, S. Skakun, A. Shelestov, M. Lavreniuk, B. Yailymov, O. Kussul // *International Archives of the Photogrammetry, Remote Sensing & Spatial Information Sciences*. — 2015. — P. 45–52.

9. Лавренюк Н. С. Классификация больших площадей земного покрова по ретроспективным спутниковым данным / Н. С. Лавренюк, С. В. Скакун, А. Ю. Шелестов, Б. Я. Яйлимов, С. Л. Янчевский, Д. Ю. Ящук, А. М. Костецкий // Кибернетика и системный анализ. — 2016. — Т. 52, № 1. — С. 137-149.
10. Куcssуль Н. М. Ретроспективна регіональна карта земного покpиву для України: методологія побудови та аналіз результатів / Н. М. Куcssуль, А. Ю. Шелестов, С. В. Скакун, Р. М. Басараб, Б. Я. Яйлимов, М. С. Лавренюк, А. В. Колотій, Д. Ю. Ящук // Космічна наука і технологія. — 2015. — Т. 21, № 3. — С. 31–39.
11. Kolotii A. Comparison of biophysical and satellite predictors for wheat yield forecasting in Ukraine / A. Kolotii, N. Kussul, A. Shelestov, S. Skakun, B. Yailymov, R. Basarab, M. Lavreniuk, T. Oliinyk, V. Ostapenko // International Archives of the Photogrammetry, Remote Sensing & Spatial Information Sciences. — 2015. — P. 39–44.
12. Скакун С. В. Класифікація сільськогосподарських посівів з використанням часових рядів супутникових даних / С. В. Скакун, А. Ю. Шелестов, Б. Я. Яйлимов, В. А. Остапенко, М. С. Лавренюк, А. В. Вікулов // Індуктивне моделювання складних систем. — 2014. — Вип. 6. — С. 157–166.
13. Шелестов А. Ю. Регресійний аналіз показників розвитку рослинництва в регіонах України за статистичними і супутниковими даними / А. Ю. Шелестов, І. М. Бутко, М. С. Лавренюк, Б. Я. Яйлимов, А. В. Колотій // Індуктивне моделювання складних систем. — 2015. — Вип. 7. — С. 282–290.
14. Куcssуль Н. Н. Анализ изменений земного покрова на основе технологии глубинного машинного обучения / Н. Н. Куcssуль, Н. С. Лавренюк, А. Ю. Шелестов, Б. Я. Яйлимов, И. Н. Бутко // Проблемы управления и информатики. — 2016. — № 3. — С. 140–151.
15. Kussul N. Land Cover Changes Analysis Based on Deep Machine Learning Technique / N. Kussul, N. Lavreniuk, A. Shelestov, B. Yailymov, I. Butko // Journal of Automation and Information Sciences. – 2016. – Vol. 48, No. 5. – P. 42-54. DOI: 10.1615/JAutomatInfScien.v48.i5.40.

16. Kussul N. Resilience Aspects in the Sensor Web Infrastructure for Natural Disaster Monitoring and Risk Assessment Based on Earth Observation Data / N. Kussul, S. Skakun, A. Shelestov, O. Kussul, B. Yailymov // IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing. — 2014. — Vol. 7, No. 9 — P. 3826–3832.
17. Шелестов А. Ю. Структурно-функциональный анализ сервис-ориентированных систем / А. Ю. Шелестов, О. М. КуССуль, Б. Я. Яйлимов // Індуктивне моделювання складних систем: Зб. наук. пр. — К.: МННЦ ІТС НАН та МОН України — 2012. — Вип. 4. — С. 224–238.
18. Яйлимов Б. Я. Інформаційна технологія оцінки збитків від посух / Б. Я. Яйлимов // Збірник матеріалів V міжнародної науково-практичної конференції студентів, аспірантів та молодих вчених "Інформаційні технології: економіка, техніка, освіта", (м. Київ, НУБіП України, 13-14 листопада 2014р.). — 2014. — С. 152–153.
19. Скакун С. В. Метод оцінки ризиків повеней на основі часових рядів супутникових даних / С. В. Скакун, Р. М. Басараб, Б. Я. Яйлимов, Р. Климчук // Збірник матеріалів V міжнародної науково-практичної конференції студентів, аспірантів та молодих вчених "Інформаційні технології: економіка, техніка, освіта", (м. Київ, НУБіП України, 13-14 листопада 2014 р.). — 2014. — С. 112–113.
20. Остапенко В. А. Метод оцінки площ сільськогосподарських культур / В. А. Остапенко, Б. Я. Яйлимов, А. В. Вікулов // Збірник матеріалів V міжнародної науково-практичної конференції студентів, аспірантів та молодих вчених "Інформаційні технології: економіка, техніка, освіта", (м. Київ, НУБіП України, 13-14 листопада 2014 р.). — 2014. — С. 142–144.
21. Скакун С. В. Інформаційна технологія оцінки ризиків затоплень / С. В. Скакун, А. В. Колотій, Б. Я. Яйлимов, Б. В. Супрун // Збірник матеріалів V міжнародної науково-практичної конференції студентів, аспірантів та молодих вчених "Інформаційні технології: економіка, техніка, освіта", (м. Київ, НУБіП України, 13-14 листопада 2014р.). — 2014. — С. 150–151.

22. Яйлимов Б. Я. Геопортал «Сич-2» в контексте GEOSS и информационной безопасности / Б. Я. Яйлимов, О. М. Куssуль, С. В. Скаун, В. Н. Крыгин // XV міжнародна молодіжна науково-практична конференція "Людина і космос". Матеріали доповідей, (м. Дніпропетровськ, 2013р.). — 2013. — С. 481.
23. Басараб Р. М. Географічна прив'язка даних космічного апарату ДЗЗ «СІЧ-2» з використанням GRID технологій / Р. М. Басараб, О. М. Кравченко, Б. Я. Яйлимов, В. М. Кригін, М. С. Лавренюк // XV міжнародна молодіжна науково-практична конференція "Людина і космос". Матеріали доповідей, (м. Дніпропетровськ, 2013р.). — 2013. — С. 470.
24. Скаун С. В. Моніторинг надзвичайних ситуацій на основі високопродуктивних GRID обчислень / С. В. Скаун, А. Ю. Шелестов, Б. Я. Яйлимов, В. М. Кригін // Обчислюваний інтелект (результат, проблеми, перспективи): Матеріали 2-ї Міжнародної науково-технічної конференції (14-17 травня 2013 р., Черкаси). — Черкаси: Маклаут. — 2013. — С. 242.
25. Шелестов А. Ю. Grid-технології як один із оперативних засобів для оновлення геопорталів / А. Ю. Шелестов, Б. Я. Яйлимов // Матеріали міжнародної науково-практичної конференції "Глобальні та регіональні проблеми інформатизації в суспільстві та природокористуванні 2013", (м. Київ, НУБіП України, 13-14 червня 2013р.). Наукове видання. — К.:ТОВ "Агро Медіа Груп". — 2013. — С. 26–27.
26. Шелестов А. Ю. Геопортал полігонних та дистанційних вимірювань проекту JESAM-Ukraine / А. Ю. Шелестов, Т. А. Баранова, Р. М. Басараб, Б. Я. Яйлимов // Матеріали міжнародної науково-практичної конференції "Глобальні та регіональні проблеми інформатизації в суспільстві та природокористуванні 2013", (м. Київ, НУБіП України, 13-14 червня 2013р.). Наукове видання. — К.:ТОВ "Агро Медіа Груп". — 2013. — С. 25.
27. Куssуль Н. М. Визначення аномалії розвитку рослинності / Н. М. Куssуль, О. М. Кравченко, С. В. Скаун, Б. Я. Яйлимов // Збірник матеріалів Міжнародної науково-практичної конференції "Глобальні та регіональні проблеми інформатизації в суспільстві та природокористуванні 2013"(13-14

червня 2013 року, Київ, Україна). — Київ: ТОВ «Аграр Медіа Груп». — 2013. — С. 27–28.

28. Куcssуль Н. М. Валідація глобальних продуктів в рамках проекту JESAM / Н. М. Куcssуль, А. Ю. Шелестов, С. В. Скаун, О. М. Кравченко, А. І. Міронов, Б. Я. Яйлимов, Р. М. Басараб, В. А. Остапенко // Наукове видання. 13-а Українська конференція з космічних досліджень, (2-6 вересня 2013р., м. Євпаторія, Україна). — Видавництво "Кафедра". — 2013. — С. 164.

29. Куcssуль Н. М. Класифікація літніх сільськогосподарських культур на основі злиття різнорідних наземних та супутникових даних / Н. М. Куcssуль, А. Ю. Шелестов, Б. Я. Яйлимов // Наукове видання. 13-а Українська конференція з космічних досліджень, (2-6 вересня 2013р., м. Євпаторія, Україна). — Видавництво «Кафедра». — 2013. — С. 181.

30. Куcssуль Н. М. JESAM project in Ukraine: current state / Н. М. Куcssуль, А. Ю. Шелестов, С. В. Скаун, О. М. Кравченко, А. В. Колотій, Б. Я. Яйлимов // Збірник матеріалів 13-ї української конференції з космічних досліджень (2-6 вересня 2013 року, Євпаторія, Крим, Україна). — 2013. — С. 173.

31. Шелестов А. Ю. Автоматизация выявления аномалий развития вегетации растительности по спутниковым данным / А. Ю. Шелестов, Б. Я. Яйлимов // Сборник тезисов докладов конференции. Одиннадцатая Всероссийская открытая конференция "Современные проблемы дистанционного зондирования Земли из космоса", Москва, Россия. — 2013.

32. Басараб Р. М. Grid-сервіс побудови маски затоплених територій на основі радіолокаційних даних / Р. М. Басараб, Б. Я. Яйлимов, С. Г. Скачков // Аерокосмічні спостереження в інтересах сталого розвитку та безпеки: Матеріали доповідей Четвертої Міжнародної конференції "GEO-UA 2014" (2014 р., м. Київ). — ISBN 978-966-02-7248-4 (електронне видання). — 2014. — С. 159–160.

33. Agarwal R. A new Bayesian ensemble of trees approach for land cover classification of satellite imagery / R. Agarwal, P. Ranjan, H. Chipman // Canadian Journal of Remote Sensing. — 2014. — Vol. 39, No. 6. — P. 507-520.

34. Басараб Р. М. Grid-сервіс розв'язання задачі параметризації гідрологічної моделі / Р. М. Басараб, Б. Я. Яйлимов, А. А. Ткаченко // Аерокосмічні спостереження в інтересах сталого розвитку та безпеки: Матеріали доповідей Четвертої Міжнародної конференції "GEO-UA 2014" (2014 р., м. Київ). — ISBN 978-966-02-7248-4 (електронне видання). — 2014. — С. 161–162.
35. Басараб Р. М. Класифікація пошкоджень сільськогосподарських культур Красногвардійського району АР Крим / Р. М. Басараб, Б. Я. Яйлимов, Д. Ю. Ящук, М. Ю. Приймак // Аерокосмічні спостереження в інтересах сталого розвитку та безпеки: Матеріали доповідей Четвертої Міжнародної конференції "GEO-UA 2014" (2014 р., м. Київ). — ISBN 978-966-02-7248-4 (електронне видання). — 2014. — С. 163–164.
36. Куусуль Н. М. Геопортал українського сегменту JESAM / Н. М. Куусуль, А. Ю. Шелестов, С. В. Скакун, В. А. Остапенко, В. М. Кригін, К. В. Рева, Б. Я. Яйлимов, А. І. Міронов // Аерокосмічні спостереження в інтересах сталого розвитку та безпеки: Матеріали доповідей Четвертої Міжнародної конференції "GEO-UA 2014" (2014 р., м. Київ). — ISBN 978-966-02-7248-4 (електронне видання). — 2014. — С. 203–204.
37. Куусуль Н. М. Макет геопорталу українського сегменту системи МАКСМ / Н. М. Куусуль, А. Ю. Шелестов, С. В. Скакун, В. А. Остапенко, В. М. Кригін, К. В. Рева, Б. Я. Яйлимов, А. І. Міронов // Аерокосмічні спостереження в інтересах сталого розвитку та безпеки: Матеріали доповідей Четвертої Міжнародної конференції "GEO-UA 2014" (2014 р., м. Київ). — ISBN 978-966-02-7248-4 (електронне видання). — 2014. — С. 205–206.
38. Куусуль Н. М. Моніторинг та оцінка збитків від посух на основі супутникових даних / Н. М. Куусуль, А. Ю. Шелестов, С. В. Скакун, Б. Я. Яйлимов // Аерокосмічні спостереження в інтересах сталого розвитку та безпеки: Матеріали доповідей Четвертої Міжнародної конференції "GEO-UA 2014" (2014 р., м. Київ). — ISBN 978-966-02-7248-4 (електронне видання). — 2014. — С. 209–210.
39. Шелестов А. Ю. Верифікація глобальних продуктів на основі на основі

наземних досліджень в рамках проекту JESAM / А. Ю. Шелестов, Н. М. Куцуль, С. В. Скакун, А. І. Міронов, В. А. Остапенко, Б. Я. Яйлимов // Аерокосмічні спостереження в інтересах сталого розвитку та безпеки: Матеріали доповідей Четвертої Міжнародної конференції “GEO-UA 2014” (2014 р., м. Київ). — ISBN 978-966-02-7248-4 (електронне видання). — 2014. — С. 246–247.

40. Шелестов А. Ю. Grid-технології в задачах моніторингу надзвичайних ситуацій / А. Ю. Шелестов, С. В. Скакун, А. В. Колотій, Б. Я. Яйлимов // Аерокосмічні спостереження в інтересах сталого розвитку та безпеки: Матеріали доповідей Четвертої Міжнародної конференції “GEO-UA 2014” (2014 р., м. Київ). — ISBN 978-966-02-7248-4 (електронне видання). — 2014. — С. 248–249.

41. Шелестов А. Ю. Анализ эффективности сервис-ориентированной системы на базе структурно-функционального анализа / А. Ю. Шелестов, Б. Я. Яйлимов // Всеукраїнська наукова конференція «Сучасні проблеми математичного моделювання та обчислювальних методів», (22 – 23 лютого 2013 року, м. Рівне, РДГУ - НУВГП). — 2013. — С. 201.

42. Куцуль Н. М. Картографування земного покриву для великих територій / Н. М. Куцуль, М. С. Лавренюк, А. Ю. Шелестов, Б. Я. Яйлимов // Наукове видання. 14-а Українська конференція з космічних досліджень, (8-12 вересня 2014 р., м. Ужгород). — 2014. — С. 174.

43. Шелестов А. Ю. Грідифікація задачі картографування повеней за радіолокаційними даними / А. Ю. Шелестов, Р. М. Басараб, Б. Я. Яйлимов, Скачков С. // Матеріали II Міжнародної науково-практичної конференції "Глобальні та регіональні проблеми інформатизації в суспільстві та природокористуванні 2014", (м. Київ, НУБіП України, 26-27 червня 2014 р.). — К.: Аграр Медіа Груп. — 2014. — С. 11.

44. Скакун С. Метод оцінювання ризику посух на основі часових рядів супутникових даних та теорії екстремальних величин / С. Скакун, Р. Басараб, Б. Яйлимов, Д. Ящук, Р. Климчук // Матеріали II Міжнародної науково-

практичної конференції "Глобальні та регіональні проблеми інформатизації в суспільстві та природокористуванні 2014", (м. Київ, НУБіП України, 26-27 червня 2014 р.). — К.: Аграр Медіа Груп. — 2014. — С. 12–14.

45. Басараб Р. Задача підбору параметрів моделі LISFLOOD-FP з використанням технології Грід / Басараб Р., Яйлимов Б., Баранова Т., Ткаченко А. // Матеріали II Міжнародної науково-практичної конференції "Глобальні та регіональні проблеми інформатизації в суспільстві та природокористуванні 2014", (м. Київ, НУБіП України, 26-27 червня 2014 р.). — К.: Аграр Медіа Груп, 2014 р. — С. 14.

46. Шелестов А. Ю. Метод злиття геопросторових даних в задачі картографування земної поверхні / А. Ю. Шелестов, Б. Я. Яйлимов // Матеріали Міжнародної наукової конференції «Сучасні проблеми математичного моделювання та обчислювальних методів» (м. Рівне, 19 – 22 лютого 2015 року). — Рівне : РВВ РДГУ. — 2015. — С. 180.

47. Лавренюк М. С. Порівняння різних методів для побудови маски сільськогосподарських земель / М. С. Лавренюк, С. В. Скакун, Б. Я. Яйлимов // Матеріали Міжнародної наукової конференції «Астрономічна школа молодих вчених. Актуальні проблеми астрономії і космонавтики» (м. Житомир, 20 – 22 травня 2015 року). — Київ-Житомир. — 2015. — С. 51.

48. Яйлимов Б. Я. Метод злиття геопросторових даних в задачах побудови карт сільськогосподарських посівів / Б. Я. Яйлимов, М. С. Лавренюк // Матеріали Міжнародної наукової конференції «Астрономічна школа молодих вчених. Актуальні проблеми астрономії і космонавтики» (м. Житомир, 20 – 22 травня 2015 року). — Київ-Житомир. — 2015. — С. 89.

49. Яйлимов Б. Методи злиття даних на основі прийняття рішень для побудови карти класифікації / Б. Яйлимов, М. Лавренюк // Матеріали III Міжнародної науково-практичної конференції «Глобальні та регіональні проблеми інформатизації в суспільстві та природокористуванні 2015», (м. Київ, НУБіП України, 25-26 червня 2015 р.). — К.: Інтерсервис. — 2015 р. — С. 20.

50. Шелестов А. Експеримент SPOT-5 TAKE 5 для тестового полігону JESAM в

Україні / А. Шелестов, Б. Яйлимов, О. М. Костецький, Р. Басараб, А. Колотій // Матеріали III Міжнародної науково-практичної конференції "Глобальні та регіональні проблеми інформатизації в суспільстві та природокористуванні 2015", (м. Київ, НУБіП України, 25-26 червня 2015 р.). — К.: Інтерсервіс. — 2015 р. — С. 18.

51. Лавренюк М. Класифікація сільськогосподарських посівів в межах міжнародного експерименту JESAM / М. Лавренюк, С. Скакун, Б. Яйлимов // Матеріали III Міжнародної науково-практичної конференції "Глобальні та регіональні проблеми інформатизації в суспільстві та природокористуванні 2015", (м. Київ, НУБіП України, 25-26 червня 2015 р.). — К.: Інтерсервіс. — 2015 р. — С. 15.

52. Костецький О. Побудова карт біофізичних параметрів з використанням даних наземних вимірювань та супутникових продуктів в проєкті IMAGINES / О. Костецький, А. Шелестов, А. Колотій, Р. Басараб, Б. Яйлимов, В. Трохименко, Г. Загородня // Матеріали III Міжнародної науково-практичної конференції "Глобальні та регіональні проблеми інформатизації в суспільстві та природокористуванні 2015", (м. Київ, НУБіП України, 25-26 червня 2015 р.). — К.: Інтерсервіс. — 2015 р. — С. 10.

53. Lavreniuk M. Regional Retrospective High Resolution Land Cover For Ukraine: Methodology And Results / Lavreniuk M., Kussul N., Skakun S., Shelestov A., Yailymov B. // International Geoscience and Remote Sensing Symposium 2015 (IGARSS 2015), № 15599383, — P. 3965-3968. DOI: 10.1109/IGARSS.2015.7326693.

54. Яйлимов Б. Я. Методи злиття даних в задачі картографування земного покриття / Б. Я. Яйлимов // Матеріали XIV Всеукраїнської науково-практичної конференції студентів, аспірантів та молодих вчених «Теоретичні і прикладні проблеми фізики, математики та інформатики», (26-28 травня 2016 р., м. Київ, Україна). — 2016. — Т. 2. — С. 184-185.

55. Яйлимов Б. Я. Злиття супутникових даних на основі ансамблевого методу класифікації / Б. Я. Яйлимов, М. С. Лавренюк // Materials of reports of the 5th

International Conference “GEO-UA 2016” (October 10-14, 2016, Kyiv). — 2016. — P. 60 – 61.

56. Giria C. Next generation of global land cover characterization, mapping, and monitoring / B. Pengrab, J. Longc, T. R. Lovelanda // Int. J. Appl. Earth Observation and Geoinform. — 2013. — Vol. 25. — P. 30—37.

57. Land Cover (GLCNMO) - Global version // [Електронний ресурс] Режим доступу : <http://www.iscgm.org/gm/glcnm.html>.

58. Зацерковний В. І. Геоінформаційні системи і бази даних: монографія. / В. І. Зацерковний, В. Г. Бурачек, О. О Железняк, А. О. Терещенко. // Ніжин : НДУ ім. М. Гоголя — 2014. — Кн. 1/3-38. — 492 с.

59. Чандра А. М. Дистанционное зондирование и географические информационные системы / А. М. Чандра, С. К. Гош. - М.: Техносфера. — 2008. — 312 с.

60. World Programme for the Census of Agriculture // [Електронний ресурс] Режим доступу : <http://www.fao.org/economic/ess/ess-wca/en/>.

61. Gong X. Land Use Classification / X.Gong, L.Gunnar, S.Tsuji // [Електронний ресурс] / Режим доступу : http://unstats.un.org/unsd/envaccounting/londongroup/meeting14/LG14_10a.pdf.

62. Coordinating land use / land cover analyses and data in Europe // [Електронний ресурс] Режим доступу : http://cordis.europa.eu/project/rcn/39991_en.pdf.

63. Development of a harmonised framework for multi-purpose land cover / land use information systems derived from earth observation data // [Електронний ресурс] Режим доступу : http://cordis.europa.eu/project/rcn/39999_en.pdf.

64. Groom G. Strategic landscape monitoring for the Nordic countries / G. Groom, T. Reed // Copenhagen : Nordic Council of Ministers. – 2001. – 129 p.

65. Gregorio A. Land Cover Classification System (LCCS) / Antonio Di Gregorio, Louisa J.M. Jansen // [Електронний ресурс] FAO Режим доступу : http://www.fao.org/docrep/003/x0596e/X0596e01e.htm#P213_18188.

66. Africover Eastern Africa module (2013) // [Електронний ресурс] Режим доступу : http://www.glcen.org/activities/africover_en.jsp

67. Classification for Land Use Statistics: Eurostat Remote Sensing Programme (1993) [Электронный ресурс] EUROSTAT Режим доступа : http://ec.europa.eu/eurostat/ramon/nomenclatures/index.cfm?TargetUrl=LST_NOM_DTL&StrNom=CLUSTERS93&StrLanguageCode=EN&IntPcKey=&StrLayoutCode=HIERARCHIC.
68. LUCAS, 2009. Land Use and Cover Area frame Survey. // [Электронный ресурс] Режим доступа : http://ec.europa.eu/eurostat/documents/205002/208938/LUCAS2009_C1-Instructions_Revised20130925.pdf/.
69. Martino L. Land Use/Cover Area frame statistical Survey: Methodology and Tools / L. Martino, M. Fritz // Statistics in focus. – 2008. – No. 33 : Eurostat. - Luxembourg. – P.1-7.
70. CORINE Land Cover. [Электронный ресурс] / Режим доступа : <http://www.eea.europa.eu/publications/COR0-landcover>.
71. CORINE Land cover - Part 1: Methodology [Электронный ресурс] European Environment Agency Режим доступа : <http://www.eea.europa.eu/publications/COR0-part1>.
72. IMAGE2000 and CLC2000 Products and Methods [Электронный ресурс] Institute for Environment and Sustainability Режим доступа : http://image2000.jrc.ec.europa.eu/reports/image2000_products_and_methods.pdf.
73. Friedl M. A. Global land cover mapping from MODIS: algorithms and early results / M. A. Friedl, D. K. McIver, J.C.F. Hodges, X.Y. Zhang, D. Muchoney, A.H. Strahler, C.E. Woodcock, S. Gopal, A. Schneider, A. Cooper, A. Baccini, F. Gao, C. Schaaf // Remote Sensing of Environment. — 2002. — Vol. 83, No. (2-Jan). — P. 287–302.
74. Copernicus Land Monitoring Services. [Электронный ресурс] / Режим доступа: <http://land.copernicus.eu/pan-european/corine-land-cover>.
75. Loveland T. R. Development of a global land cover characteristics database and IGBP DISCover from 1 km AVHRR data / T. R. Loveland, B. C. Reed, J. F. Brown, D. O. Ohlen, Z. Zhu, L. Yang, J. W. Merchant // International Journal of Remote Sensing. – 2000. – Vol. 21. – No. 6-7. – P. 1303-1330.

76. Bontemps S. GLOBCOVER 2009-Products description and validation report / S. Bontemps, P. Defourny, E. Van Bogaert // ESA. – 2010. – P. 1-30.
77. Roy D.P. Landsat-8: Science and product vision for terrestrial global change research / D.P. Roy, M. A. Wulder, T. R. Loveland, et al. // Remote Sensing of Environment. — 2014. — Vol. 145. — P. 154–172.
78. Conrad C. Per-Field Irrigated Crop Classification in Arid Central Asia Using SPOT and ASTER Data / C. Conrad, S. Fritsch, J. Zeidler, G. Rücker, S. Dech // Remote Sens. – 2010. – Vol. 2, No. 4. – P. 1035–1056. DOI:10.3390/rs2041035
79. Peña-Barragán J.M. Object-based crop identification using multiple vegetation indices, textural features and crop phenology / J.M. Peña-Barragán, M.K. Ngugi, R.E. Plant, J. Six, // Remote Sensing of Environment. —2011. — Vol. 115, No. 6. — P. 1301–1316.
80. Yang C. Evaluating high resolution SPOT 5 satellite imagery for crop identification / C. Yang, J.H. Everitt, D. Murden // Computers and Electronics in Agriculture. — 2011. — Vol. 75, No. 2. — P. 347–354.
81. Геопортал ESA CCI Land Cover products: a new generation of satellite-derived global land cover products. Электронный ресурс — режим доступа станом на 15.12.2014: <http://maps.elie.ucl.ac.be/CCI/viewer/index.php>.
82. Chen J. Global land cover mapping at 30m resolution: A POK-based operational approach / J. Chen, A. Liao, X. Cao, et al. // ISPRS J. Photogrammetry and Remote Sens. — 2015. — Vol. 103. — P. 7—27.
83. Boryan C. Monitoring US agriculture: the US Department of Agriculture, National Agricultural Statistics Service, cropland datalayer program / C. Boryan, Z. Yang, R. Mueller, M. Craig // Geocarto International. — 2011. — Vol. 26, No. 5. — P. 341–358.
84. Yu L. FROM-GC: 30 m global cropland extent derived through multisource data integration / L. Yu, J. Wang, N. Clinton, Q. Xin, L. Zhong, Y. Chen, P. Gong // International Journal of Digital Earth. — 2013. — Vol. 6, No. 6. — P. 521–533.
85. Yu L. Improving 30 m global landcover map FROM-GLC with time series MODIS and auxiliary data sets: a segmentation-based approach / L. Yu, J. Wang, P.

Gong // International Journal of Remote Sensing. — 2013. — Vol. 34, No. 16. — P. 5851-5867.

86. McNairn H. Integration of optical and Synthetic Aperture Radar (SAR) imagery for delivering operational annual crop inventories / H. McNairn, C. Champagne, J. Shang, D. A. Holmstrom, G. Reichert // ISPRS Journal of Photogrammetry and Remote Sensing. — 2009. — Vol. 64, No. 5. — P. 434–449.

87. Coppin P. Digital change detection methods in ecosystem monitoring: a review / P. Coppin, I. Jonckheere, K. Nackaerts, B. Muys // Int. J. Remote Sens. — 2004. — Vol. 25, No. 9. — P. 1565–1596.

88. Gong P. Assessment of GLC map accuracies using flux net location data // Progr. Nat. Sci. — 2009. — Vol. 19. — P. 754-759.

89. Fritz S. Comparison of global and regional land cover maps with statistical information for the agricultural domain in Africa / S. Fritz, L. See, F. Rembold // Int. J. Remote Sens. — 2010. — Vol. 31, No. 9. — P. 2237–2256.

90. Jia Kun Land cover classification using Landsat 8 operational land imager data in Beijing, China / Kun Jia et al. // Geocarto International. — 2014. — Vol. 29, No 8. — P. 941-951.

91. Gallego F. J. Crop area estimates through remote sensing: stability of the regression correction / F. J. Gallego, J. Delincé, C. Rueda // International Journal of Remote Sensing. — 1993. — Vol. 14, No.18. — P. 3433-3445.

92. Zhong L. Efficient corn and soybean mapping with temporal extendability: a multi-year experiment using landsat imagery / L. Zhong, P. Gong, G. S. Biging // Remote Sensing Environ. — 2014. — Vol. 140. — P. 1–13.

93. Atzberger C. Obtaining crop-specific time profiles of NDVI: the use of unmixing approaches for serving the continuity between SPOT-VGT and PROBA-V time series / C. Atzberger, A. R. Formaggio, Y. E. Shimabukuro, T. Udelhoven, M. Mattiuzzi, G. A. Sanchez, & E. Arai // International Journal of Remote Sensing. — 2014. — Vol. 35, No. 7. — P. 2615-2638.

94. Malenovský Z. Sentinels for science: Potential of Sentinel-1,-2, and-3 missions for scientific observations of ocean, cryosphere, and land / Z. Malenovský, H. Rott, J.

Cihlar, M. E. Schaepman, G. García-Santos, R. Fernandes, & M. Berger // *Remote Sensing of Environment*. — 2012. — Vol. 120. — P. 91-101.

95. Барталев С. А. Дистанционная оценка параметров сельскохозяйственных земель по спутниковым данным спектрорадиометра MODIS / С. А. Барталев, Е. А. Лупян, И. А. Нейштадт, И. Ю. Савин // *Современные проблемы дистанционного зондирования Земли из космоса: Физические основы, методы и технологии мониторинга окружающей среды. Сборник научных статей*. — М.: «GRAND polygraph». — 2005. — Том 2. — С. 228-236.

96. Повх В. И. Оценка структуры посевов подсолнечника в Ставропольском крае по данным дистанционного зондирования спектрорадиометра MODIS / В. И. Повх, Г. П. Гарбузов, Л. А. Шляхова // *Современные проблемы дистанционного зондирования Земли из космоса*. — 2007. — Том 2. — Вып. 4. — С. 284-288.

97. González F. Comparing Two Methodologies for Crop Area Estimation in Spain Using Landsat TM Images and Ground Gathered Data / F. González, S. López, J. M. Cuevas // *Remote Sensing of Environment*. — 1991. — Vol. 32. — P. 29-36.

98. Hiederer R. Estimating European crop surfaces from SPOT and Landsat-TM data / R. Hiederer, J.C. Favard, D. Guedes, M. Sharman // *Proceedings of «International Symposium on Operationalization of Remote Sensing»*, ITC Enschede, The Netherlands. — 1993. — Vol. 2. — P. 116-127.

99. Taylor J. C. Regional Crop Inventories in Europe Assisted by Remote Sensing / J. C. Taylor, C. Sannier, J. Delincé, F. J. Gallego // *Synthesis Report, Office for Publications of the EC. Luxembourg*. — 1997. — P.71.

100. Taylor T. W. Agricultural analysis for a worldwide crop assessment / T. W. Taylor // *Proceedings SPOT Conference, Paris, France*. — 1996. — P. 485-488.

101. Battese G. E. An error-components model for prediction of county crop areas using survey and satellite data / G. E. Battese, R. M. Harter, W. A. Fuller // *Journal of the American Statistical Association*. — 1988. — Vol. 83. — P. 28-36.

102. Arvor D. Classification of MODIS EVI time series for crop mapping in the state of mato grosso, Brazil / D. Arvor, M. Jonathan, M. S. P. Meirelles, V. Dubreuil, L.

- Durieux // *Int. J. Remote Sensing*. — 2011. — Vol. 32, No. 22. — P. 7847–7871.
103. Abou EL-Magd I. Improvements in land use mapping for irrigated agriculture from satellite sensor data using a multi-stage maximum likelihood classification / I. Abou EL-Magd, T. Tanton // *Int. J. Remote Sensing*. — 2003. — Vol. 24, No. 21. — P. 4197-4206.
104. Blanzieri E. Nearest neighbor classification of remote sensing images with the maximal margin principle / E. Blanzieri, F. Melgani // *Geoscience and Remote Sensing, IEEE Transactions on*. — 2008. — Vol. 46, No. 6. — P. 1804-1811.
105. Cheng Q. Logistic regression for feature selection and soft classification of remote sensing data / Q. Cheng, P. K. Varshney, M. K. Arora // *Geoscience and Remote Sensing Letters, IEEE*. — 2006. — Vol. 3, No. 4. — P. 491-494.
106. Duro D. C. A comparison of pixel-based and objectbased image analysis with selected machine learning algorithms for the classification of agricultural landscapes using spot-5 hrg imagery / D. C. Duro, S. E. Franklin, M. G. Dubé // *Remote Sensing Environ*. — 2012. — Vol. 118. — P. 259–272.
107. Long, J. A. Objectoriented crop classification using multitemporal ETM+ SLC-off imagery and random forest / J. A. Long, R. L. Lawrence, M. C. Greenwood, L. Marshall, P. R. Miller // *GIScience & Remote Sensing*. — 2013. — Vol. 50, No. 4. — P. 418–436.
108. Simon Haykin. *Neural Networks and Learning Machines* (3rd Edition), // Prentice Hall, Upper Saddle River, New Jersey.— 2009.— P. 937.
109. Cortes C. Support-vector networks / C. Cortes, V. Vapnik // *Machine Learning*. — 1995. — Vol. 20, No. 3. — P. 273–297. DOI:10.1007/BF00994018
110. Yang C. Airborne hyperspectral imagery and linear spectral unmixing for mapping variation in crop yield / C. Yang, J. H. Everitt, J. M. Bradford // *Precision Agr*. — 2007. — Vol. 8, No. 6. — P. 279–296.
111. Omkar S. Crop classification using biologically-inspired techniques with high resolution satellite image / S. Omkar, J. Senthilnath, D. Mudigere, M. M. Kumar // *J. Indian Soc. Remote Sensing*. — 2008. — Vol. 36, No. 2. — P. 175–182.
112. Duro D. C. A comparison of pixel-based and object based image analysis with

selected machine learning algorithms for the classification of agricultural landscapes using spot-5 hrg imagery / D. C. Duro, S. E. Franklin, M. G. Dubé // *Remote Sensing Environ.* — 2012. — Vol. 118. — P. 259–272.

113. Akbari M. Crop and land cover classification in Iran using landsat 7 imagery / M. Akbari, A. R. Mamanpoush, A. Gieske, M. Miranzadeh, M. Torabi, H. Salemi // *Int. J. Remote Sensing.* — 2006. — Vol. 27, No. 19. — P. 4117–4135.

114. Potgieter A. B. Estimating crop area using seasonal time series of enhanced vegetation index from modis satellite imagery / A. B. Potgieter, A. Apan, P. Dunn, G. Hammer // *Crop Pasture Sci.* — 2007. — Vol. 58, No. 4. — P. 316–325.

115. Mingwei Z. Crop discrimination in northern china with double cropping systems using fourier analysis of time-series modis data / Z. Mingwei, Z. Qingbo, C. Zhongxin, L. Jia, Z. Yong, C. Chongfa // *Int. J. Appl. Earth Observation Geoinform.* — 2008. — Vol. 10, No. 4. — P. 476–485.

116. Waldner F. Land cover and crop type classification along the season based on biophysical variables retrieved from multi-sensor high-resolution time series / F. Waldner, M. J. Lambert, W. Li, M. Weiss, V. Demarez, D. Morin, C. Marais-Sicre, O. Hagolle, F. Baret, P. Defourny // *Remote Sensing.* — 2015. — Vol. 7, No. 8. — P. 10400–10424.

117. Löw F. Impact of feature selection on the accuracy and spatial uncertainty of per-field crop classification using support vector machines / F. Löw, U. Michel, S. Dech, C. Conrad // *ISPRS J. Photogr. Remote Sensing.* — 2013. — Vol. 85. — P. 102–119.

118. De Fries R.S. Global land cover classifications at 8 km spatial resolution: the use of training data derived from Landsat imagery in decision tree classifiers / De Fries R.S., Hansen M., Townshend J.R.G., Sohlberg R. // *International Journal of Remote Sensing*, 1998. — 19. — P. 3141–3168.

119. Hansen M.C. Global land cover classification at 1 km spatial resolution using a classification tree approach / Hansen M.C., Defries R.S., Townshend J.R.G., Sohlberg R. // *International Journal of Remote Sensing*, 2000. — 21. — P. 1331–1364.

120. Xu M. Decision tree regression for soft classification of remote sensing data / M.

Xu, P. Watanachaturaporn, P. K., Varshney, M. K. Arora // Remote Sensing of Environment. – 2005. – Vol. 97, №. 3. – P. 322-336.

121. Waldner F. Annual Cropland Mapping over Medium to Large Field Agrosystems with Moderate Resolution Satellite Time-series / F. Waldner, D. Abelleira, S. Veron, M. Zhang, B. Wu, D. Plotnikov, S. Bartalev, M. Lavreniuk, S. Skakun, N. Kussul, G. Le Maire, S. Dupuy, I. Jarvis, P. Defourny // International Journal of Remote Sensing. — 2015. — P. 1-32 (accepted).

122. Барсегян А. А. Анализ данных и процессов: учеб. пособие / А. А. Барсегян, М. С. Куприянов, И. И. Холод, М. Д. Тесс, С. И. Елизаров. — 3-е изд., перераб. и доп. — СПб.: БХВ-Петербург, 2009. — 512 с.

123. Cortes C. Support-vector network / C. Cortes, V. Vapnik // Machine Learning. — 1995. — Vol. 20. — P. 1-25.

124. Ghoggali N. A multiobjective genetic SVM approach for classification problems with limited training samples / N. Ghoggali, M. Farid, B. Yakoub // IEEE Transactions on geoscience and remote sensing. – 2009. – Vol. 47, No. 6. – P. 1707-1718. DOI: 10.1109/TGRS.2008.2007128.

125. Галлего Х. Анализ эффективности различных подходов для классификации посевов на основе спутниковой и наземной информации / Х. Галлего, А. Кравченко, Н. Куусуль, С. Скакун, А. Шелестов, Ю. Грипич // Проблемы управления и информатики. — 2012. — № 3. — С. 123–134.

126. Куусуль Н. Н. Регресійні моделі оцінки врожайності культур по даним MODIS / А. Н. Кравченко, С. В. Скакун, Т. И. Адаменко, А. Ю. Шелестов, А. В. Колотий, Ю. А. Грипич // Сборник научных статей «Современные проблемы дистанционного зондирования Земли из космоса». — 2012. — Том. 9, №1. — С. 95-107.

127. Bishop C. Pattern Recognition and Machine Learning / Bishop C.// Springer, 2006. — 740 p.

128. Архангельский В.И. Нейронные сети в системах автоматизации / В.И. Архангельский, И.Н. Богаенко, Г.Г. Грабовский, Н.А. Рюмшин // – К.: Техніка, 1999. – 364 с.

129. Kussul N. Neural network approach for user activity monitoring in computer networks / N. Kussul, S. Skakun // IEEE International Conference on Neural Networks - Conference Proceedings. — 2004. — Vol. 2. — P. 1557-1561.
130. Skakun S.V. Analysis of Applicability of Neural Networks for Classification of Satellite Data / S.V. Skakun, E.V. Nasuro, A.N. Lavrenyuk, O.M. Kussul // Journal of Automation and Information Sciences. — 2007. — Vol. 39, No. 3. — P. 37–50. DOI: 10.1615/JAutomatInfScien.v39.i3.40.
131. Kussul E.M. Adaptive neural network classifier with multfloat input coding / E.M. Kussul, T.N. Baidyk, V.V. Lukovich, and D.A. Rachkovskij // Proc. of 6th Int. Conf. “Neuro-Nimes 93”, 25-29 Oct. 1993, Nimes, France. — 1993. — P. 209–216.
132. Kussul E.M. Multilevel assembly neural architecture and processing of sequences / E.M. Kussul, D.A. Rachkovskij // Neurocomputers and Attention. — 1991. — Vol. 39. — P. 577–590.
133. Gallego J. Efficiency Assessment of Different Approaches to Crop Classification Based on Satellite and Ground Observations / J. Gallego, A. Kravchenko, N. Kussul, S. Skakun, A. Shelestov, Y. Grypych // Journal of Automation and Information Sciences. — 2012. — Vol. 44, No. 5. — P. 67-80.
134. Kussul N. Grid and sensor web technologies for environmental monitoring / N. Kussul, A. Shelestov, S. Skakun // Earth Science Informatics. — 2009. — Vol. 2, No. 1-2. — P. 37-51.
135. Kussul N. Interoperable Infrastructure for Flood Monitoring: SensorWeb, Grid and Cloud / N. Kussul, D. Mandl, K. Moe, J.P. Mund, J. Post, A. Shelestov, S. Skakun, J. Szarzynski, G. Van Langenhove, M. Handy // IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing. — 2012. — Vol. 5, No. 6. — P. 1740–1745.
136. Shelestov A.Yu. Using the fuzzy-ellipsoid method for robust estimation of the state of a grid system node / A.Yu. Shelestov, N.N. Kussul // Cybernetics and Systems Analysis. — 2008. — Vol. 44, No. 6. — P. 847-854.
137. Kussul N. Determination of Inundated Territories on the Basis of Integration of Heterogeneous Data / N. Kussul, E. Lupian, A. Shelestov, S. Skakun, Yu.

- Tishchenko, L. Hluchy // *Journal of Automation and Information Sciences*. – 2007. – Vol. 39, No. 9. – P. 42–51. DOI: 10.1615/JAutomatInfScien.v39.i12.40.
138. Kussul N. The Wide Area Grid Testbed for Flood Monitoring Using Earth Observation Data / N. Kussul, A. Shelestov, S. Skakun, G. Li, O. Kussul // *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*. – 2012. – Vol. 5, No. 6. – P. 1746-1751.
139. Shelestov A. Grid technologies in monitoring systems based on satellite data / A. Shelestov, N. Kussul, S. Skakun // *Journal of Automation and Information Sciences*. — 2006. — Vol. 38, No. 3. — P. 69-80.
140. Skakun S., Kussul N. An agent approach for providing security in distributed systems / S. Skakun, N. Kussul // *Modern Problems of Radio Engineering, Telecommunications and Computer Science Proceedings of International Conference, TCSET*. – 2006. – P. 212-215.
141. Kussul N. Grid technologies for satellite data processing and management within international disaster monitoring projects / N. Kussul, A. Shelestov, S. Skakun // *Grid and Cloud Database Management*. – 2011. – P. 279–305.
142. Stathaki T. *Image Fusion: Algorithms and Applications* / T. Stathaki. – Academic Press is an imprint of Elsevier. – 2008. – 492 p.
143. Chen Z. Fusion of hyperspectral and multispectral images: A novel framework based on generalization of pan-sharpening methods / Z. Chen, H. Pu, B. Wang, and G.-M. Jiang // *IEEE Geosci. Remote Sens. Lett.* – 2014. – Vol. 11, No. 8, P. 1418-1422.
144. Aiazzi B. Quality assessment of Pansharpening methods and products / B. Aiazzi, L. Alparone, S. Baronti, and A. Garzelli // *IEEE Geosci. Remote Sens. Soc. Newslett.* – 2011. – Vol. 1, No. 161– P. 10-18.
145. Kussul N. Service-oriented infrastructure for flood mapping using optical and SAR satellite data / N. Kussul, A. Shelestov, S. Skakun, G. Li, O. Kussul, J. Xie., 2014. // *International Journal of Digital Earth* . — 2014. — Vol. 7, No. 10 — P. 829-845.
146. Kussul N. Probabilistic estimation of trust model and threat resistance analysis

in service-oriented systems / N. Kussul, O. Kussul, S. Skakun // Int. J. on Information Models and Analyses. – 2012. – No 1. – P. 28-46.

147. Dikaiakos M. D. Cloud Computing: Distributed Internet Computing for IT and Scientific Research / M. D. Dikaiakos, D. Katsaros, P. Mehra, G. Pallis, A. Vakali // Internet Computing. — 2009. — Vol.13, No. 5. — P. 10-13.

148. Елманова Е. Коротко о вычислениях в «облаке» / Е. Елманова // КомпьютерПресс. — 2010. — №3. — С.102-104.

149. Черняк Л. SaaS — конец начала. // Открытые системы [Электронный ресурс].— Режим доступа: <http://www.osp.ru/os/2007/10/4706040/>.

150. Kussul N. The data fusion grid infrastructure: Project objectives and achievements / N. Kussul, A. Shelestov, S. Skakun, O. Kravchenko, Y. Gripich, L. Hluchý, P. Kopp, E. Lupian // Computing and Informatics. — 2010. — Vol. 29, No. 2. — P. 319-334.

151. Michael P. Service-Oriented Computing: State of the Art and Research Challenges / Michael P. Papazoglou, Paolo Traverso, Schahram Dustdar [et al.] // IEEE Computer. — 2007. — V.40, No.11. — P. 38-45.

152. Alonso G. Web Services: Concepts, Architectures and Applications / Gustavo Alonso. – Berlin Heidelberg: Springer-Verlag. —2004. —359 p.

153. Erl T. Service-oriented architecture: a field guide to integrating XML and Web services / Thomas Erl. — Prentice Hall PTR. — 2004. — 536 p.

154. Hurwitz J. Service Oriented Architecture For Dummies / J.Hurwitz, C.Baroudi, R.Bloor and M.Kaufman. — UK: Wiley & Sons. — 2006. — 384 p.

155. Bakan G.M. Fuzzy ellipsoidal filtering algorithm of static object state / G.M. Bakan, N.N. Kussul // Problemy Upravleniya I Informatiki (Avtomatika). — 1996. — No. 5. — P. 77-92.

156. Jagannadham D. Java2 distributed application development (socket, rmi, servlet, corba) approaches, xml-rpc and web services functional analysis and performance comparison / D. Jagannadham, V. Ramachandran, H. Kumar, N. Harish // In International Symposium on Communications and Information Technologies, Sydney, Australia, March 2007. — 2007. — P. 1337–1342.

157. Dong J. High-Assurance Service-Oriented Architectures / J. Dong, A. Raymond Paul, Liang-Jie Zhang // IEEE Computer. — 2008. — Vol. 41, No. 8. — P. 27-28.
158. Cormen T. H. Introduction to Algorithms / T. H. Cormen, C. E. Leiserson, R. L. Rivest, C. Stein. — Cambridge, Massachusetts : The MIT Press, 3rd edition, 2009. — 1312 p.
159. Tuia D. Active learning methods for remote sensing image classification / D. Tuia, F. Ratle, F. Pacifici, M. F. Kanevski, W. J. Emery // IEEE Transactions on Geoscience and Remote Sensing. — 2009. — Vol. 47, No. 7. — P. 2218–2232.
160. Kussul N. N. Neural networks learning using method of fuzzy ellipsoidal estimates / N. N. Kussul // J. Autom. Inf. Sci. — 2001. — Vol. 33, No. 3. — P. 52-57.
161. Skakun S. V. Reconstruction of Missing Data in Time-Series of Optical Satellite Images Using Self-Organizing Kohonen Maps / S. V. Skakun, R. M. Basarab // Journal of Automation and Information Sciences. — 2014. — Vol. 46, No. 12. — P. 19–26.
162. Kussul N. Parcel based classification for agricultural mapping and monitoring using multi-temporal satellite image sequences / N. Kussul, G. Lemoine, J. Gallego, S. Skakun, M. Lavreniuk // The International Geoscience and Remote Sensing Symposium (IGARSS), 2015 IEEE International. — 2015. — P. 165-168. — DOI: 10.1109/IGARSS.2015.7325725.
163. Kussul N. The use of satellite SAR imagery to crop classification in Ukraine within JECAM project / Kussul N., Skakun S., Shelestov A., Kussul O. // IEEE International Geoscience and Remote Sensing Symposium (IGARSS). — 2014. — P. 1497–1500.
164. Warmerdam F. The geospatial data abstraction library / Eds H. Brent, L. G. Michael // Open Source Approaches in Spatial Data Handling. — Berlin: Springer, 2008. — Vol. 2. — P. 87-104.
165. Sim J. The Kappa Statistic in Reliability Studies: Use, Interpretation, and Sample Size Requirements / J. Sim, C. C Wright // Physical Therapy Journal. — 2005. — Vol. 85, No. 3. — P. 257–268.
166. Skakun S. Restoration of Missing Data due to Clouds on Optical Satellite

Imagery Using Neural Networks / S. Skakun, N. Kussul, R. Basarab // ESA Sentinel-2 for Science Workshop, ESA-ESRIN, Frascati, Rome, Italy, 20-22 May 2014. — 2014. Available at: http://seom.esa.int/S2forScience2014/files/05_S2forScience-MethodsI_KUSSUL.pdf.

167. Skakun S. Dealing with missing data in remote sensing images within land and crop classification / S. Skakun, N. Kussul, R. Basarab // 40th COSPAR Scientific Assembly. Held 2-10 August 2014, in Moscow, Russia, Abstract A3.1-3-14.

168. Скакун С.В. Восстановление отсутствующих данных во временных рядах оптических спутниковых снимков с использованием самоорганизующихся карт Кохонена / С.В. Скакун, Р.М. Басараб // Проблеми управління та інформатики — 2014., №6. — С. 88 - 94.

169. Hansen M. C. Continuous fields of land cover for the conterminous United States using Landsat data: first results from the Web-Enabled Landsat Data (WELD) project / M. C. Hansen, A. Egorov, D. P. Roy, P. Potapov, J. Ju, S. Turubanova, I. Kommareddy, T. R. Loveland // Remote Sensing Letters. — 2011. — Vol. 2, No. 4. — P. 279-288.

170. Büttner G. The CORINE land cover 2000 project / G. Büttner, J. Feranec, G. Jaffrain, L. Mari, G. Maucha, T. Soukup // EARSeL eProceedings. — 2004. — Vol.3, No. 3. — P. 331–346.

171. Olofsson P. Good practices for estimating area and assessing accuracy of land change / P. Olofsson, G. M. Foody, M. Herold, S. V. Stehman, C. E. Woodcock, M. A. Wulder // Remote Sensing of Environment. — 2014. — Vol. 148. — P. 42–57.

172. Foody G. M. Status of land cover Classification accuracy assessment / G. M. Foody // Remote Sensing of Environment. — 2002. — Vol. 80, No. 1. — P. 185–201.

173. Skakun S. Validation of Global EO Biophysical Products at JECAM Test Site in Ukraine / S. Skakun, N. Kussul, O. Kravchenko, R. Basarab, V. Ostapenko, B. Yailymov, A. Shelestov, A. Kolotii, A. Mironov // 40th COSPAR Scientific Assembly. 2-10 August 2014, in Moscow, Russia, Abstract A3.1-59-14.

174. Kussul N. Application of satellite optical and SAR images for crop mapping and

- area estimation in Ukraine / N. Kussul, S. Skakun, O. Kravchenko, A. Shelestov, F. J. Gallego, O. Kussul // International Journal "Information Technologies & Knowledge". — 2013. — Vol. 7, No. 3 — P. 203–210.
175. Shelestov A. JECAM Activities in Ukraine / A. Shelestov, S. Skakun, R. Basarab, T. Baranova // EARTH Bioresources and Life Quality. — 2013. — Vol. 4. — P. 1–13.
176. Shelestov A.Yu. Geospatial information system for agricultural monitoring / A.Yu. Shelestov, A.N. Kravchenko, S.V. Skakun, S.V. Voloshin, N.N. Kussul // Cybernetics and Systems Analysis. — 2013. — Vol. 49, No. 1. — P. 124–132.
177. Lavreniuk M. Integrated Use of Multi-temporal SAR and Optical Satellite Imagery for Crop Mapping in Ukraine / M. Lavreniuk, N. Kussul, S. Skakun // American Geophysical Union Fall Meeting (AGU 2014), San Francisco, USA, December 15–19, 2014.
178. Kussul N. Large Scale Crop Classification in Ukraine using Multi-temporal Landsat-8 Images with Missing Data / N. Kussul, S. Skakun, A. Shelestov, M. Lavreniuk // American Geophysical Union Fall Meeting (AGU 2014), San Francisco, USA, December 15–19, 2014.
179. Геопортал Геологічної служби США (United States Geological Survey) // [Електронний ресурс] Режим доступу: <http://earthexplorer.usgs.gov>.
180. Chander G. Summary of current radiometric calibration coefficients for Landsat MSS, TM, ETM+, and EO-1 ALI sensors / G. Chander, B. Markham, D. Helder // Remote Sens. Environ. — 2009. — Vol. 113, No. 5. — P. 893–903.
181. Gong P. Finer resolution observation and monitoring of global land cover: first mapping results with Landsat TM and ETM+ data / P. Gong, J. Wang, L. Yu, Y. Zhao, L. Liang, ... & C. Li // International Journal of Remote Sensing. — 2013. — Vol. 34, No. 7. — P. 2607–2654.
182. Skakun S. Flood Hazard and Flood Risk Assessment Using a Time Series of Satellite Images: A Case Study in Map_classification / S. Skakun, N. Kussul, A. Shelestov, O. Kussul // Risk Analysis. — 2014. — Vol. 34, No. 8. — P. 1521–1537. DOI: 10.1111/risa.12156.

183. Kussul N.N. Disaster risk assessment based on heterogeneous geospatial information / N.N. Kussul, B.V. Sokolov, Ya.I. Zyelyk, S.V. Skakun, A.Yu. Shelestov // *Journal of Automation and Information Sciences*. — 2010. — Vol. 42, No. 12. — P. 32–45.
184. Павлов А.А. Информационные технологии и алгоритмизация в управлении / А.А.Павлов, С.Ф.Теленик. — К.: Техника. — 2002. С. — 344 с.
185. Lecca G. Grid computing technology for hydrological applications / G. Lecca, M. Petitdidier, L. Hluchy, M. Ivanovic, N. Kussul, N. Ray, V. Thieron // *Journal of Hydrology*. — 2011. — Vol. 403, No. 1-2. — P. 186-199.
186. Згуровський М. З. Системний аналіз: проблеми, методологія, застосування / М. З. Згуровський, Н. Д. Панкратова. // К. : Наук. Думка. — 2005. — 744 с.
187. КуССуль Н.Н. Grid-системы для задач исследования Земли. Архитектура, модели и технологии / Н.Н. КуССуль, А.Ю. Шелестов // — К.: Наукова думка. — 2008. — 452 с.
188. Мордвинов В. И. Численные методы анализа и прогноза погоды : учеб. пособие / В. И. Мордвинов. — М. : Изд-во Иркут. гос. ун-та. — 2008. — 143 с.
189. Kussul O. Assessing security threat scenarios for utility-based reputation model in Grids / O. Kussul, N. Kussul, S. Skakun // *Computers & Security*. — 2013. — Vol. 34, P. 1-15.
190. Kravchenko A. Water resource quality monitoring using heterogeneous data and high-performance computations / A. Kravchenko, N. Kussul, E. Lupian, V. Savorsky, L. Hluchy, A. Shelestov // *Cybernetics and Systems Analysis*. — 2008. — Vol. 44, No. 4. — P. 616-624. DOI:10.1007/s10559-008-9032-x
191. Kussul N. High-performance intelligent computations for environmental and disaster monitoring / N. Kussul, A. Shelestov, S. Skakun, O. Kravchenko // *Int. J. Information Technologies & Knowledge*. — 2009. — Vol. 3. — P. 135–156.
192. Kussul N. Technologies for Satellite Data Processing and Management Within International Disaster Monitoring Projects / N. Kussul, A. Shelestov, S. Skakun // In . *Grid and Cloud Database Management Grid* — Fiore, S.; Aloisio, G. (Eds.). — 2011, Springer — P. 279–305.

193. Ji L. Assessing vegetation response to drought in the northern Great Plains using vegetation and drought indices / L. Ji, A. J. Peters // *Remote Sensing of Environment*. – 2003. – Vol. 87, No. 1. – P. 85–98.
194. Getachew B. Using Satellite Images for Drought Monitoring: A Knowledge Discovery Approach / B. Getachew, H. Shawndra, T. Tsegaye, S. Atnaфу // *Journal of Strategic Innovation and Sustainability*. – 2011. – Vol. 7. – P. 135–153.
195. Zhang X. Drought-induced vegetation stress in southwestern North America / X. Zhang, M. Goldberg, D. Tarpley, M. Friedl, J. Morisette, F. Kogan, Y. Yu // *Environmental Research*. – 2010. – Vol. 5, No. 2. – P. 511–538.
196. Щербенко Е.В. Мониторинг засухи по данным космических съемок / Е.В Щербенко // *Современные проблемы дистанционного зондирования Земли из космоса: сб. ст. М: ООО «Азбука»*. – 2007. – Т.2. – С. 395–407.
197. Bayarjargal Y. A comparative study of NOAA-AVHRR derived drought indices using change vector analysis / Y. Bayarjargal, A. Karnieli, M. Bayasgalan, S. Khudulmur, C. Gandush, C. Tucker // *Remote Sens. Environ.* – 2006. – Vol. 105. – P. 9–22.
198. Kussul N. Crop area estimation in Ukraine using satellite data within the MARS project 2012 / N. Kussul, S. Skakun, A. Shelestov, O. Kravchenko, J. F. Gallego, O. Kussul // *IEEE International Geoscience and Remote Sensing Symposium (IGARSS)*. — 2012. — P. 3756–3759.
199. Mitchell H. B. *Multi-sensor Data Fusion* / H. B. Mitchell // Springer-Verlag, Berlin. – ISBN 9783540714637.
200. Kussul N. Crop state and area estimation in Ukraine based on remote and in-situ observations / N. Kussul, A. Shelestov, S. Skakun, O. Kravchenko, B. Moloshnii // *Int. J. on Information Models and Analyses*. — 2012. — Vol. 1, No. 3. — P. 251–259.
201. Kussul N. Flood Monitoring from SAR Data / N. Kussul, A. Shelestov, S. Skakun // *NATO Science for Peace and Security Series C: Environmental Security*. — 2011. — P. 19–29. DOI: 10.1007/978-90-481-9618-0_3.
202. Плющиков В.Г. Методы управления сельскохозяйственными рисками:

учебно-методическое пособие / В.Г. Плющиков, Э.А. Довлетярова, Н.И. Ильясова // Гос. образоват. учреждение высш. проф. образования "Рос. ун-т дружбы народов". – Москва: Изд-во Российского ун-та дружбы народов. – 2006. – С. 70.

203. Memarsadeghi N. Fast Implementation of the ISODATA Clustering Algorithm / N. Memarsadeghi, N. Netan-yahu, J. LeMoignc // International Journal of Computational Geometry and Applications. – 2007. – Vol. 17. – P. 71–103.

204. Настанова гідрометеорологічним станціям і постам. – Вип. 11. Агromетеорологічні спостереження. – К. : Державна гідрометеорологічна служба України. – 2007. – 303 с.

205. Tommi Jaakkola, course materials for 6.867 Machine Learning, Fall 2006. MIT OpenCourseWare (<http://ocw.mit.edu/>), Massachusetts Institute of Technology.

206. Центр «АГРО ПЕРСПЕКТИВА». Информационно-аналитический центр по вопросам аграрного и химического бизнесов // Средние закупочные цены по Украине на зерновые. 2012. [Электрон. ресурс] Режим доступа: <http://www.agroperspectiva.com/ru/price/?d=:20-08-2012>.

207. Главное управление статистики в Автономной Республике Крым. [Электронный ресурс] Режим доступа: <http://www.sf.ukrstat.gov.ua/2013Zexp45.pdf>.

ДОДАТКИ



УКРАЇНА

**МІНІСТЕРСТВО ОСВІТИ І НАУКИ УКРАЇНИ
НАЦІОНАЛЬНИЙ ТЕХНІЧНИЙ УНІВЕРСИТЕТ УКРАЇНИ
«КИЇВСЬКИЙ ПОЛІТЕХНІЧНИЙ ІНСТИТУТ
імені ІГОРЯ СІКОРСЬКОГО»**

ФІЗИКО-ТЕХНІЧНИЙ ІНСТИТУТ

03056, м. Київ, пр-т Перемоги, 37; тел/факс (+38 044) 236-70-98, e-mail: novi@pti.kpi.ua, <http://www.kpi.ua>

ДОВІДКА

про використання результатів дисертаційної роботи

Яйлимова Богдана Ялкаповича

**«Автоматизована інформаційна технологія картографування земного
покриву на основі методів та моделей злиття супутникових даних»,
поданої на здобуття наукового ступеня кандидата технічних наук
за спеціальністю 05.07.12 — «Дистанційні аерокосмічні дослідження»**

Результати дисертаційної роботи Яйлимова Б.Я. «Автоматизована інформаційна технологія картографування земного покриву на основі методів та моделей злиття супутникових даних» використовуються в навчальному процесі фізико-технічного інституту НТУУ «КПІ імені Ігоря Сікорського», зокрема, в курсах «Методи штучного інтелекту» та «Спеціальні розділи програмування».

Розроблені здобувачем методи злиття різнорідних геопросторових даних (векторних та супутникових) використано при підготовці циклу лабораторних робіт з курсу «Аналіз геопросторових даних 1. Методи аналізу великих геопросторових даних» та застосовуються студентами Фізико-технічного інституту НТУУ «КПІ імені Ігоря Сікорського» при виконанні наукових досліджень.

Довідку видано для представлення у спеціалізовану вчену раду Інституту космічних досліджень НАН України та Державного космічного агентства України Д 26.205.01.

Т.в.о. директора ФТІ
к. ф.-м.н



С.А. Смирнов



УКРАЇНА

**ДЕРЖАВНА СЛУЖБА УКРАЇНИ З НАДЗВИЧАЙНИХ СИТУАЦІЙ
УКРАЇНСЬКИЙ ГІДРОМЕТЕОРОЛОГІЧНИЙ ЦЕНТР**

01601, м. Київ-30, вул. Золотоворітська, 6Б, тел/факс. 279-1080, 239-9387, E-mail:
office@meteo.gov.ua

Довідка

Про використання результатів дисертаційної роботи
Яйлимова Богдана Ялкаповича
«Автоматизована інформаційна технологія картографування земного покриття
на основі методів та моделей злиття супутникових даних»,
поданої на здобуття наукового ступеня кандидата технічних наук
за спеціальністю 05.07.12 — «Дистанційні аерокосмічні дослідження»

Результати дисертаційної роботи Яйлимова Б.Я., а саме карти класифікації сільськогосподарських культур побудовані із використанням розроблених методів злиття супутникових даних (Landsat, Sentinel) для території всієї України з просторовим розрізненням 10-30 м, є об'єктивним джерелом інформації про структуру посівних площ та можуть використовуватись в УкрГМЦ ДСНС України. Зокрема карти посівів, що отримані за участі в проекті Sen2Agri, є першим продуктом національного рівня, отриманим з просторовим розрізненням 10 м. Подібні карти можуть використовуватись для оцінки площ та завчасного прогнозування валового збору сільськогосподарських культур.

Довідку видано для представлення у спеціалізовану вчену раду Інституту космічних досліджень НАН України та Державного космічного агентства України Д 26.205.01.

Начальник відділу агрометеорології
УкрГМЦ ДСНС України,
кандидат географічних наук



Адаменко Т.І.



НАЦІОНАЛЬНИЙ УНІВЕРСИТЕТ
БІОРЕСУРСІВ І ПРИРОДОКОРИСТУВАННЯ УКРАЇНИ

ФАКУЛЬТЕТ ІНФОРМАЦІЙНИХ ТЕХНОЛОГІЙ

03041, м. Київ, вул. Героїв Оборони, 16, навчальний корпус № 15, тел./факс: (044) 527-83-52,
E-mail: o-glazunova@nubip.edu.ua

29.10.2016 № 1836
На № _____ від _____

Довідка

Про використання результатів дисертаційної роботи
Яйлимова Богдана Ялкаповича
**«Автоматизована інформаційна технологія картографування земного
покриву на основі методів та моделей злиття супутникових даних»,**
поданої на здобуття наукового ступеня кандидата технічних наук
за спеціальністю 05.07.12 — «Дистанційні аерокосмічні дослідження»

Результати дисертаційної роботи Яйлимова Б.Я., а саме методи класифікації супутникових даних та сервіс-орієнтована автоматизована інформаційна технологія картографування земного покриву, використано для побудови карти літніх та озимих культур в межах виконання науково-дослідної теми № 509-пр «Створення геоінформаційних сервісів підтримки прийняття рішень в сфері АПК для органів державної та місцевої влади», що виконується в Національному університеті біоресурсів і природокористування України у 2015-2016 рр.

Декан факультету інформаційних технологій
Національного університету біоресурсів і
природокористування України, д.пед.н., доц.



О.Глазунова