

МІНІСТЕРСТВО ОСВІТИ І НАУКИ УКРАЇНИ
НАЦІОНАЛЬНИЙ АВІАЦІЙНИЙ УНІВЕРСИТЕТ
ФАКУЛЬТЕТ НАЗЕМНИХ СПОРУД І АЕРОДРОМІВ
КАФЕДРА АЕРОКОСМІЧНОЇ ГЕОДЕЗІЇ ТА ЗЕМЛЕУСТРОЮ

ДОПУСТИТИ ДО ЗАХИСТУ

Завідувач кафедри

_____Юрій ВЕЛИКОДСЬКИЙ

«__» _____ 2024 р.

**КВАЛІФІКАЦІЙНА РОБОТА
(ПОЯСНЮВАЛЬНА ЗАПИСКА)**

ЗДОБУВАЧА ОСВІТНЬОГО СТУПЕНЯ «БАКАЛАВР»

Тема: «Геоінформаційне картографування наслідків руйнації Каховської ГЕС за даними дистанційного зондування»

Виконавець: Головач Вікторія Вікторівна, студентка групи ГС-412Б

Керівник: Беленок Вадим Юрійович, кандидат фізико-математичних наук,
доцент _____

Нормоконтролер: Іщенко Наталія Федорівна, PhD, доцент _____

Київ 2024

НАЦІОНАЛЬНИЙ АВІАЦІЙНИЙ УНІВЕРСИТЕТ

Факультет наземних споруд і аеродромів

Кафедра аерокосмічної геодезії та землеустрою

Спеціальність 193 «Геодезія та землеустрій»

Освітньо-професійна програма «Геоінформаційні системи і технології»

ЗАТВЕРДЖУЮ

Завідувач кафедри

_____Юрій ВЕЛИКОДСЬКИЙ

«__» _____ 2024 р.

ЗАВДАННЯ

на виконання кваліфікаційної роботи

Головач Вікторії Вікторівні

1. Тема кваліфікаційної роботи: «Геоінформаційне картографування наслідків руйнації Каховської ГЕС за даними дистанційного зондування», затверджена наказом ректора від «22» квітня 2024р. № 601/ст.
2. Термін виконання роботи: з «20» травня 2024 року по «16» червня 2024 року.
3. Вихідні дані роботи: космічні знімки, топографічні карти, підручники, публікації та наукові статті за тематикою роботи.
4. Зміст пояснювальної записки: у першому розділі розглянуто погляди інших дослідників, охарактеризовано об'єкт дослідження. У другому розділі охарактеризовано методи, алгоритми та дані, які використані під час класифікації земного покриву досліджуваної території. У третьому розділі наведені отримані результати, їх аналіз та інтерпретація, порівняння сучасних даних і за минуле століття.
5. Перелік обов'язкового ілюстративного матеріалу: 6 таблиць, 70 рисунків.
6. Календарний план-графік

№ з/п	Завдання	Термін виконання	Підпис керівника
1	Отримано завдання. Пошуки теоретичних матеріалів	20.05-23.05.2024	
2	Опрацювання наукових джерел, формування змісту	24.05-26.05.2024	
3	Робота над першим розділом. Підготовка ілюстративного матеріалу	27.05-30.05.2024	
4	Робота над другим розділом. Визначення методів, алгоритмів та даних дослідження	31.05-04.06.2024	
5	Робота над третім розділом. Підготовка даних, опрацювання практичної частини, підготовка ілюстративного матеріалу	05.06-09.06.2024	
6	Формування висновків, оформлення списку використаних джерел	10.06-11.06.2024	
7	Підготовка роботи до захисту. Захист роботи	12.06-16.06.2024	

Дата видачі завдання: «20» травня 2024 р.

Керівник кваліфікаційної роботи: _____ Беленок В.Ю.

Завдання прийняла до виконання: _____ Головач В.В.

РЕФЕРАТ

Кваліфікаційна робота на тему: «Геоінформаційне картографування наслідків руйнації Каховської ГЕС за даними дистанційного зондування» містить: 85 сторінок, 70 рисунків, 6 таблиць, 29 використаних джерел.

Об'єктом дослідження є наслідки руйнації Каховської ГЕС.

Мета роботи: оцінка наслідків руйнації Каховської ГЕС за допомогою геоінформаційних технологій та даних дистанційного зондування для створення картографічних матеріалів.

Методи дослідження: наукові - вивчення вже існуючих даних, аналіз літературних джерел; спостереження – спостереження за досліджуваними територіями; вимірювання – знаходження площ затоплених ділянок; порівняння – для виявлення розбіжностей між періодами до та після підриву греблі; аналіз і синтез – для отриманих результатів; узагальнення - підведення підсумків.

Результатом бакалаврської роботи є створені картографічні матеріали наслідків руйнації Каховської ГЕС.

КАХОВСЬКА ГЕС, КАРТОГРАФУВАННЯ, ДАНІ ДИСТАНЦІЙНОГО ЗОНДУВАННЯ, КЛАСИФІКАЦІЯ, ЗЕМЛЕКОРИСТУВАННЯ.

ЗМІСТ

ПЕРЕЛІК УМОВНИХ СКОРОЧЕНЬ	6
ВСТУП	7
РОЗДІЛ 1. АНАЛІЗ СТАНУ ДОСЛІДЖЕНЬ	9
1.1. Характеристика Каховської ГЕС	9
1.2. Пошкодження та підрив дамби.....	12
1.3. Наслідки підриву греблі для України	16
РОЗДІЛ 2. ДАНІ ДИСТАНЦІЙНОГО ЗОНДУВАННЯ ЯК ОСНОВА ДОСЛІДЖЕННЯ	22
2.1. Поняття про метод навчання Random Forest для класифікації	22
2.2. Алгоритми процесу класифікації в GEE.....	32
2.3. Дані як основа дослідження	40
РОЗДІЛ 3. АНАЛІЗ ОТРИМАНИХ РЕЗУЛЬТАТІВ	49
3.1. Класифікація знімків методом RF в GEE	49
3.2. Виявлення затоплених територій за допомогою радарних даних	62
3.3. Аналіз змін типів землекористувань	65
ВИСНОВКИ.....	72
СПИСОК ВИКОРИСТАНИХ ДЖЕРЕЛ.....	74
ДОДАТКИ.....	78

ПЕРЕЛІК УМОВНИХ СКОРОЧЕНЬ

BAEI - Built-up Area Extraction Index;
BCI - Biophysical Composition Index;
GEE – Google Earth Engine;
IBI - Index-based Built-up Index;
LULC – land use land cover;
NDSI - Normalized Difference Snow Index;
NDVI - Normalized Difference Vegetation Index;
NDWI - Normalized Difference Water Index;
RF – Random Forest;
ROI – region of interest;
UI – Urban Index;
ГЕС – гідроелектростанція.

ВСТУП

Актуальність теми. Водосховища в Україні відіграють важливу роль, насамперед для задоволення потреб населення, адже водні ресурси з кожним роком все більш зменшуються, особливо на півдні країни. Саме ці проблеми дозволило вирішити Каховське водосховище, а саме водопостачання та забезпечення постійним й надійним зрошенням сільськогосподарських земель південних областей. Завдяки йому також було побудовано Північно-Кримський канал, що вирішило завдання водопостачання прісної води до Кримського півострова. Таким чином, Каховське водосховище стало надважливим об'єктом для півдня України.

Руйнація греблі Каховської ГЕС 6 червня 2023 року призвела до знищення водосховища, що в свою чергу, негативно вплинуло не тільки на території, які було затоплено водою, а й взагалі на прилеглі області. Постраждали як люди, так і рослинний та тваринний світи. Це спричинило знищення цілих екосистем, порушення інфраструктури, погіршення стану аграрного сектору економіки та загибелі місцевого населення.

Так як, в Україні триває війна і лівий берег Херсонської області окупований, немає можливості вільно і точно провести дослідження територій, які постраждали від даної катастрофи, і визначити масштаби трагедії для регіону. Саме для вирішення цього питання можна використати дані дистанційного зондування, які можуть надати актуальні дані на той період, що потрібен, оцінити і дослідити пошкодженні чи зруйновані території, а також провести певний аналіз для подальшої стратегії мінімізації негативних наслідків.

Метою даного дослідження є оцінка наслідків руйнації Каховської ГЕС за допомогою геоінформаційних технологій та даних дистанційного зондування для створення картографічних матеріалів.

Для досягнення поставленої мети в кваліфікаційній роботі необхідно вирішити наступні **завдання**:

1. Дослідити Каховську ГЕС та водосховище.
2. Визначити методи та підходи аналізу для виявлення наслідків.
3. Проаналізувати погляди дослідників в даній темі.
4. Виконати аналіз даних дистанційного зондування на досліджувану територію.
5. Оцінити наслідки та зробити відповідні картографічні матеріали з використанням ГІС.

Об'єктом дослідження в даній роботі є наслідки руйнації Каховської ГЕС.

Предметом дослідження даної роботи є геоінформаційне картографування наслідків за даними дистанційного зондування.

Методи дослідження: наукові - вивчення вже існуючих даних, аналіз літературних джерел; спостереження – спостереження за досліджуваними територіями; вимірювання – знаходження площ затоплених ділянок; порівняння – для виявлення розбіжностей між періодами до та після підриву греблі; аналіз і синтез – для отриманих результатів; узагальнення - підведення підсумків.

Практичне значення отриманих результатів. Дані, отримані в ході виконання кваліфікаційної роботи, можуть бути використані органами місцевого самоврядування Херсонської області, фахівцями в області захисту екології.

РОЗДІЛ 1. АНАЛІЗ СТАНУ ДОСЛІДЖЕНЬ

1.1. Характеристика Каховської ГЕС

Каховська гребля, побудована в 1956 році, була другою побудованою ГЕС на Дніпрі починаючи з кінця 1920-х рр. Її розмістили нижче течії Дніпро ГЕС і приблизно на 60 км вище течії від міста Херсона, створивши при цьому найбільше водосховище на річці Дніпро з об'ємом $18,2 \text{ км}^3$ води при нормальному рівні утримання 16,0 м над середнім рівнем моря [1].

Дана ГЕС складалася з чотирьох частин: лівобережної та правобережної ґрунтових дамб загальною довжиною 3,8 км, бетонної водозливної дамби довжиною 447 м з 28 водоскидами, великого генераторного залу з шістьма турбінами та транспортного шлюзу шириною 18 м (рис. 1.1). Максимальна висота бетонної дамби – 29 м. За проектом споруда повинна була мати пропускну спроможність $21\,400 \text{ м}^3/\text{с}$, але використовували лише – $2600 \text{ м}^3/\text{с}$, при цьому здатна виробляти 335 000 кВт електроенергії та щорічне виробництво 1,4 млрд кВт/год [2].



Рис. 1.1. Панорама будівництва Каховської ГЕС (20 серпня 1954 р.). Джерело [3]

Наповнювати Каховське водосховище почали в липні 1955 р. і закінчили навесні 1958 р. При нормальному рівні утримання 16,0 м його проектна площа

складала 2222 км² з об'ємом зберігання 22,6 км³. При рівні мертвого запасу (рівень, при якому використання води користувачами стає обмеженим) площа складала 1917 км², а залишок водойми – 11,4 км³ (рис. 1.2).

Водосховище мало другий (після Кременчуцького) корисний об'єм — 6,78 км³. Площа водної поверхні — 2155 км². Довжина по вісі становить 230 км, максимальна ширина — до 25,0 км, середня — 9,3 км. Середня глибина при нормальному підпірному рівні 8,5 м; максимальна — 24,0 м. Розрахунковий напір — 16,5 м. Мінералізація води 253–433 мг/дм³ [2; 4, с. 121].



Рис. 1.2. Вид шлюзу Каховської ГЕС (м. Нова Каховка Херсонської обл., вересень 1959 р.). Джерело [3]

Перший гідроагрегат Каховської ГЕС було введено в дію 18 жовтня 1955 р., другий — 26 листопада того ж року. У цілому на ній встановлено шість вертикальних поворотно-лопатевих гідроагрегатів, п'ять з яких останніми роками реконструйовано.

Дана споруда має кілька особливостей. На відміну від інших, вона розташована біля лівого берега. Особливістю станції є й наявність донного водоскиду. Однак його використання супроводжувалося значною вібрацією станції, через що ці отвори закрили.

Каховський гідровузол має однокамерний шлюз, що, як і ГЕС, тяжіє до лівого берега. Його довжина — 270 м, ширина — 18 м. Нині цей шлюз є найстарішим з тих, що експлуатуються, — його введено в дію 30 червня 1955 р. [4, с. 122; 5, с. 35].

Майже одночасно з будівництвом Каховської дамби почали будувати великий Північно-Кримський канал (рис. 1.3), який мав подавати воду з Дніпра до сільського господарства та Кримського півострову. Канал починається біля дамби і тягнеться майже до міста Керч у східному Криму, загальною довжиною 400,5 км.

У 1970-х рр. також був побудований Головний Каховський зрошувальний канал для забору води з водосховища, який бере початок за кілька кілометрів на північний схід від початку Північно-Кримського каналу, довжиною 129,8 км. Крім гідроенергетики, першочерговою метою створення було надійне зрошення та водопостачання півдня України [1, 6].

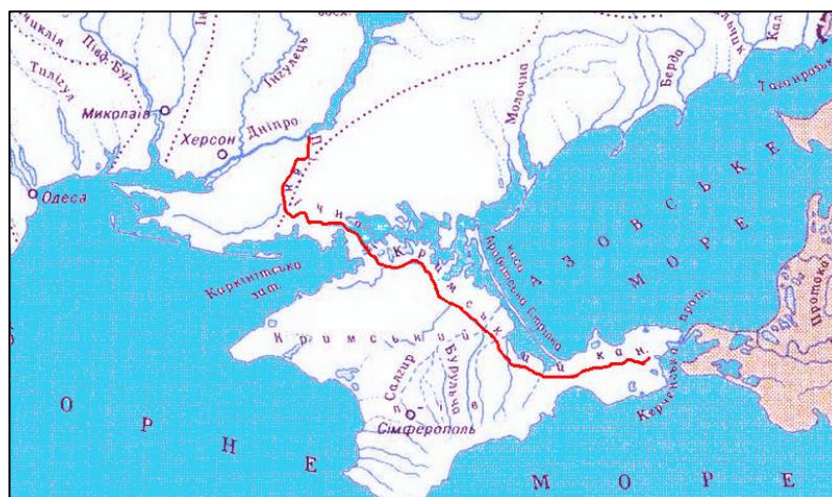


Рис. 1.3. Північно-Кримський канал

У 2020–2021 роках річні надходження води в канал досягали 0,9–1,0 км³, що забезпечує водою близько 240 000 га території. Окрім виробництва електроенергії та зрошення, воно широко використовувалося для питного та промислового водопостачання. Такі енергетичні об'єкти, як Криворізька ТЕС, Запорізька ТЕЦ, Запорізька АЕС (ЗАЕС), залежали від води для охолодження з водойми. Крім Кривого Рогу, воно забезпечувало питною водою ще кілька міст, зокрема Нікополь, Марганець і Берислав на правому березі та Енергодар, Кам'янку-Дніпровську та Дніпрорудне на лівому.

Згадані вище зрошувальні канали також забезпечували водою кілька міст і багато сіл. Було покращено умови для судноплавства, зокрема для руху суден із

зерном, рудою та металами, забезпечивши судноплавні шляхи глибиною 3,65 м, а також широко використовувалося для розведення риби та промислу [2].

1.2. Пошкодження та підрив дамби

11 листопада 2022 року російськими військовими було зруйновано три ділянки водоскиду та частину проїжджої частини через Каховську дамбу біля правого берега. Після пошкодження водоскиду та протягом наступної зими рівень води у водосховищі почав падати [2, 7].

Потрібно розуміти, що перепад висоти рівня води у водосховищі відносно русла річки нижче дамби не перевищував 13,5-14 м за звичайних умов та досягав 16 м перед підривом дамби (було зменшено пропускну здатність гідрозатворів та шлюзових камер, що призвело до збільшення рівня води до критично небезпечного рівня). При проектуванні водосховища виконувались розрахунки та моделювання на можливі небезпечні ситуації. Проте, вони не передбачали умов руйнації самого тіла дамби та її основи, обмежуючись лише умовами небезпечних паводків [8].

Глибоко всередині дамби була «ахіллесова п'ята». Оскільки її збудували за радянських часів, влада Росії мала всі інженерні креслення і знала, де вона знаходиться. Гребля була побудована з величезним бетонним блоком в основі. Через нього проходить невеликий прохід, до якого можна дістатися з машинного відділення греблі (рис. 1.4) [9].

На початку травня рівні води, виміряні в Нікополі, досягли 17,13 м над рівнем моря, це був найвищий рівень води з початку створення водосховища [1].

Перед вибухом рівень води на станції Нікополь становив 16,76 м, а загальний об'єм водосховища був надзвичайно високим – $\sim 19,8 \text{ км}^3$ (рис. 1.5).

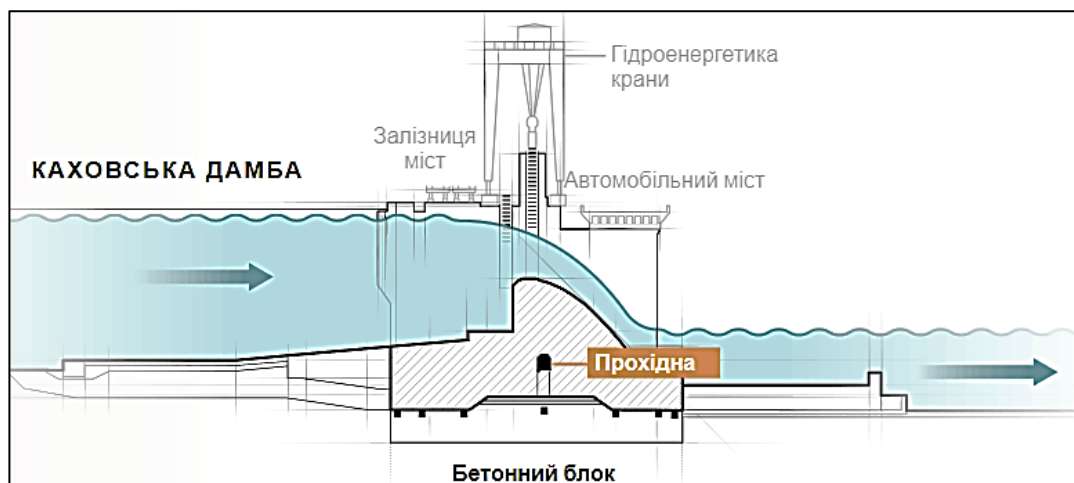


Рис. 1.4. Будова Каховської дамби (розріз)

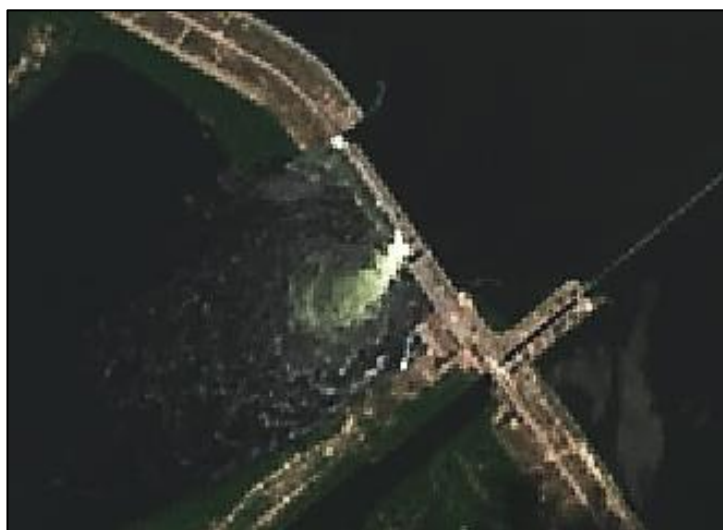


Рис. 1.5. Знімок Каховської ГЕС (дані Sentinel-2 за 03.06.2023)

У ніч на 6 червня 2023 року дамба в Каховці була зруйнована вибухом, що негайно призвело до масового витoku води з водосховища та розмивання основних частин дамби. Незабаром після вибуху рівень води в Херсоні почав підніматися. О 4:00 6 червня рівень води на вимірювальній станції в Херсоні становив 0,31 м над рівнем моря. О 8:00 він сягав 1,60 м, рівень водосховища опустився до 16,13 м. До 8:00 наступного дня він піднявся до 5,29 м, рівень водосховища впав до 14,48 м (рис. 1.6).

Максимальна позначка 5,68 м була 8 червня о 15:00. До 20:00 11 червня рівень водосховища зменшився до 9,04 м, була втрачена основна частина об'єму

водойми. Таким чином, загальний підйом рівня води в м. Херсон склав 5,37 м (рис. 1.7) [1, 2].



Рис. 1.6. Знімок зруйнованої греблі (дані Махар Technologies за 07.06.2023).

Джерело [10]

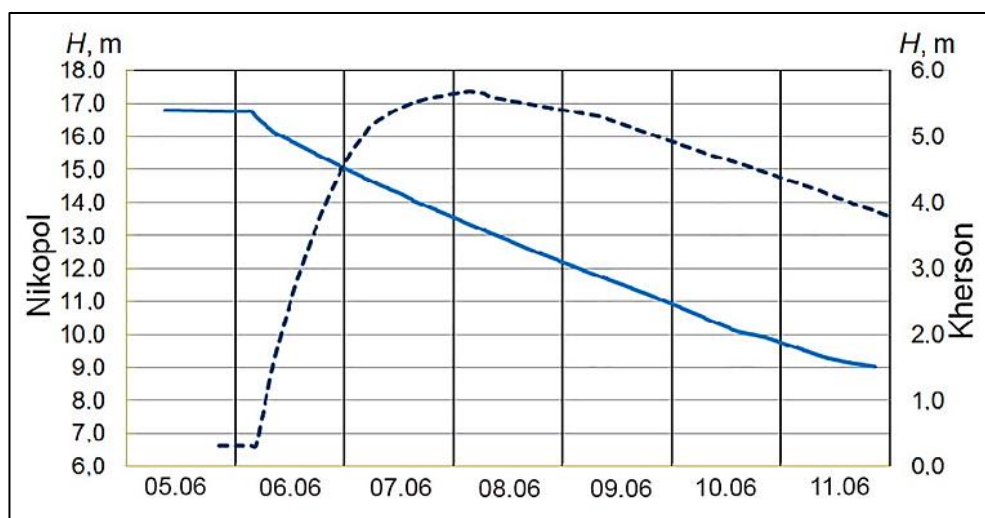


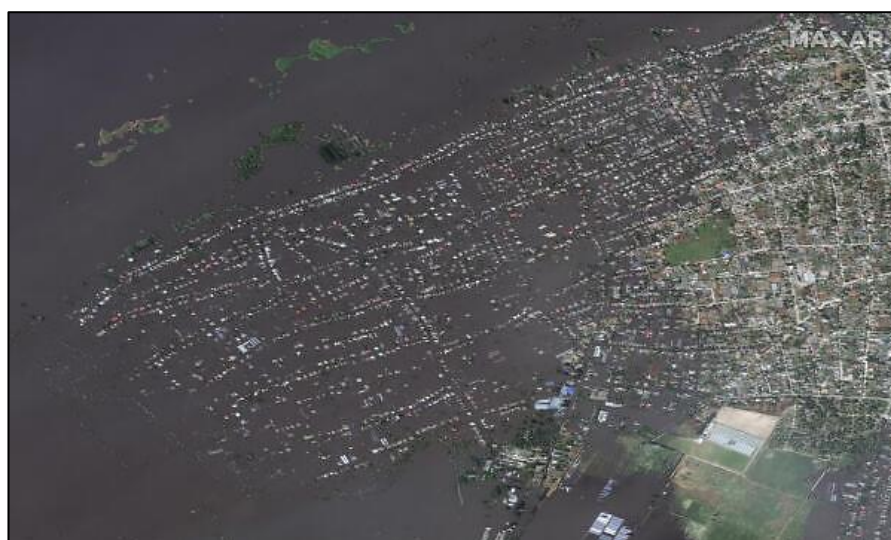
Рис. 1.7. Зміна рівня води у Херсоні (пунктир) та Нікополі (суцільна лінія)

За даними двох американських інженерів, враховуючи супутникові та сейсмічні дані про вибухи в цьому районі, найвірогіднішою причиною обвалення став заряд вибухівки, розміщений у проході для обслуговування або галереї, яка проходить через бетонне серце конструкції [11]. Вода, що вилася з водосховища, затопила близько 80 населених пунктів вздовж берегів Дніпра, з

яких частково чотири міста: Нову Каховку, Олешки (рис. 1.8) та Голу Пристань на лівому березі та Херсон на правому [1, 2].



а

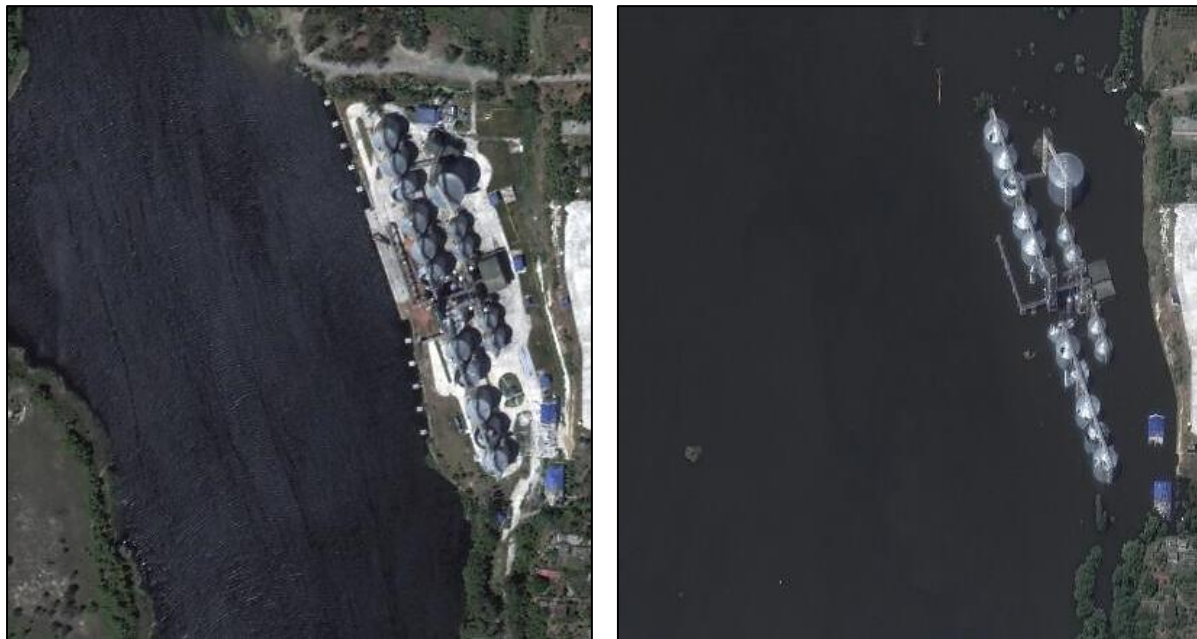


б

Рис. 1.8. Знімки міста Олешки: а – 15.05.2023, б – 07.06.2023 (дані Махаг Technologies). Джерело [10]

Підвищення рівня води спостерігалось не тільки в нижній течії р. Дніпро, а й на його правій притоці р. Інгулець, що сприяло підтопленню тут ще деяких сіл. Найбільша повінь сталася в місті Нова Каховка (рис. 1.9) та в селі Козацьке, що розташоване менш ніж за 2 км нижче за течією від зруйнованої дамби. Більшість житлових будинків у центральній частині міста Нова Каховка мають

висоту 11,5–12,0 м. З цього, за оцінками дослідників, рівень води в місті сягнув 12,5 м [1].



а

б

Рис. 1.9. Знімки міста Нова Каховка: а – 05.06.2023, б – 07.06.2023 (дані Махаг Technologies). Джерело [10]

1.3. Наслідки підриву греблі для України

За даними МВС України на 18 червня 2023 року, на правому березі Дніпра загинули або зникли безвісти близько 50 осіб. Немає інформації про людські жертви лівого берегу річки, який під контролем російських військ, але, враховуючи особливості розташування поселень, можливо, кількість жертв значно більша, ніж на правому.

Втрата водосховища також означає втрату водопостачання міських центрів і сільськогосподарських угідь, залежних від води з водосховища. Потік забруднюючих речовин у північно-західну частину Чорного моря сприяв значному цвітінню водоростей, підвищенню вимірювань важких металів і закриттю пляжів для купання поблизу Одеси.

За даними Міністерства сільського господарства України, десятки тисяч гектарів сільськогосподарських угідь були втрачені через затоплення нижче дамби, а вище за течією близько 580 000 га можуть стати непридатними для зрошення найближчим часом, що загрожує ще більшим втратам врожаю, ніж відбулися протягом минулого року [1].

Зруйновано багато зрошувальних каналів, з яких оброблялися зрошувані площі. Внаслідок цього, відбудеться заболочування осушених земель нижче Дніпра, деградація ґрунтів, загине частина деревної рослинності, яка стримувала явища берегової ерозії та поширення ярів.

Під загрозою зупинки опинилися підприємства в Нікополі та Марганці Дніпропетровської області. Розкладання риби, змив великої кількості добрив з полів, отрутохімікатів, звалищ у затоплених населених пунктах, виробництвах і заводах призвело до отруєння вод Дніпра в нижній течії [7].

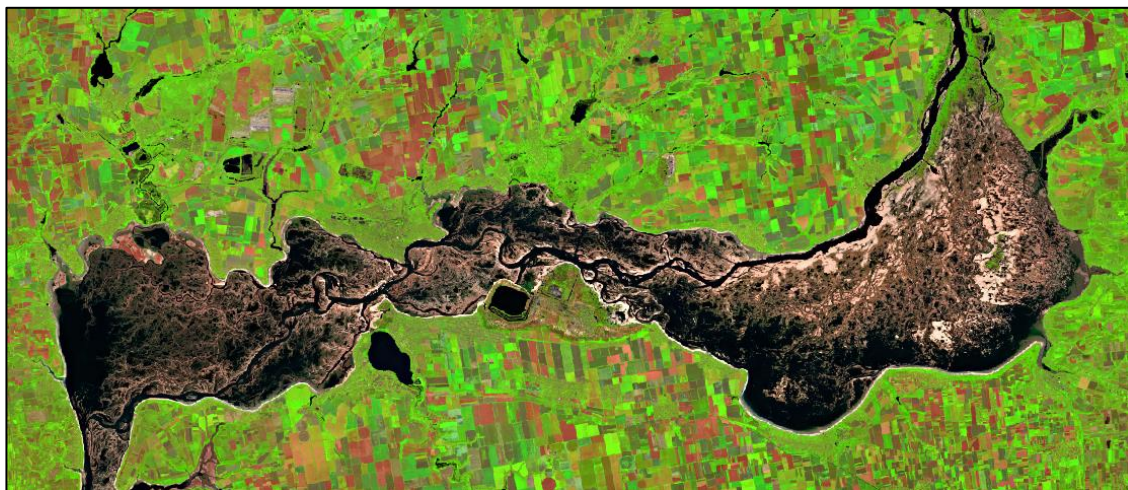
Знищення дамби також спричинило значне забруднення Дніпровсько-Бузького лиману та північно-західної частини Чорного моря. 8 червня забруднення дійшли до Одеси, а 10 – до гирла Дністра. 17 червня річкова вода, яка містила як хімічні, так і біологічні забруднення внаслідок затоплених промислових об'єктів, звалищ, каналізаційних станцій і нафтових станцій, досягла гирла річки Дунай і покрила територію площею понад 7300 км².

Спостереження також показали помітне зниження солоності моря. Середня солоність до аварії становила близько 11 г/л. Між 13 і 14 червня рівень упав до діапазону 4,2–4,4 г/л. Коли паводкові потоки почали зменшуватися, було зафіксовано відповідне збільшення солоності, що призвело до повного відновлення до кінця червня.

Приплив прісної та забрудненої води вплинув на хімічний склад морської води, і були виміряні високі рівні біогенних речовин, включаючи сполуки азоту (N) і фосфору (P). У більшості випадків найвищі концентрації спостерігалися 14 червня водночас з найменшою мінералізацією води. У морській воді біля Одеси було зафіксовано концентрацію міді (Cu) - 17,9 мкг/л, цинку (Zn) - 44,8 мкг/л і миш'яку (As) - 1,81 мкг/л, що значно перевищує контрольні рівні 0,02, 1,0 та 0,6

мкг/л відповідно. Концентрація нафти продуктів складала 0,10 мг/л, що вище за контрольний рівень 0,05 мг/л. А рівень кадмію (Cd) - 0,56 мкг/л навпаки був нижчим за контрольний стандарт у 1,5 мкг/л.

Наприкінці червня 2023 року Каховське водосховище майже повністю зникло і на місці колишнього водосховища знову виникла первісна мережа рукавів річки (рис. 1.10).



а



б

Рис. 1.10. Знімки Каховського водосховища: а – 20.06.2023, б – 05.07.2023 (дані супутника Sentinel-2)

Більша частина дна колишнього водосховища поступово вийде на поверхню. Такий ґрунт буде висихати і здебільшого покриватися тріщинами. На колишньому дні водойми з'явиться розріджена рослинність, можливо, з великою

часткою інвазивних видів. Рівень води на місці колишніх рукавів повернеться до рівня до 1956 року.

Екологічні наслідки осушення Каховського водосховища наведені у табл. 1.1, а табл. 1.2 містить в собі наслідки для підтоплених територій [11, 12].

Територія, що постраждала від цієї трагедії, становить не менше 5 000 км², які були затоплені або осушені. Близько 72% об'єму води, або 14 395 км³, було втрачено з Каховського водосховища. За підрахунками Міністерства захисту довкілля та природних ресурсів України (Міндовкілля), орієнтовна сума збитків, завданих довкіллю внаслідок вибуху дамби Каховської ГЕС, уже становить понад 55 мільярдів гривень.

Таблиця 1.1

Екологічні наслідки осушення дна Каховського водосховища та витоку води з нього

Вплив	Характеристика (пояснення)
Фауна риб	У водосховищі водиться 43 видів риб, з них 20 промислового значення. Для відновлення таких запасів знадобиться не менше 7-10 років. Знищено всі нерестовища та основний об'єм води, які є середовищем проживання риби
Земноводні і плазуни	Приблизно 15-18 видів прямо чи опосередковано постраждали від осушення водосховища. Амфібії опинилися під загрозою через руйнування місць проживання та нерестових водойм
Птахи	Через майже повне зникнення водосховища на цій території зникне кількість видів птахів, які гніздяться в цих місцях (зокрема, мартин, крячок та ін.)
Донна фауна (бентос)	Велика кількість живих організмів, що населяють воду і живуть у донному мулі вже загинули або загинуть найближчим часом. Зокрема, ряд безхребетних, насамперед молюски, різні види двокрилих, які служать кормовою базою для риб, птахів, земноводних та ін.
Рослини	Внаслідок різкого падіння рівня води, зникнуть водні та прибережно-водні рослини водосховища. Загалом оголена придонна зона водойми стане найбільшим місцем розмноження небезпечних

	інвазійних видів в регіоні
Рідкісні біотиби/типи середовищ існування	Загалом на території, що постраждала, виявлено 38 таких рідкісних видів біотопів, які перебувають під охороною Бернської конвенції. Саме для їхнього захисту в Україні були створені території Смарагдової мережі
Території природно-заповідного фонду	Постраждає низка природоохоронних територій, у тому числі щонайменше 11 об'єктів природно-заповідного фонду
Природоохоронні об'єкти міжнародного значення	Наслідки теракту матимуть негативний вплив на територію Смарагдової мережі Каховського водосховища (218 119 га) та Національного природного парку «Великий Луг» (16 755 га), водно-болотних угідь міжнародного значення архіпелагу Великі та Малі Кучугури (7740 га), заплави Сім Маяків (2140 га)

Таблиця 1.2

Екологічні наслідки підтоплення територій нижче зруйнованої дамби

Вплив	Характеристика (пояснення)
Наземна фауна	70% світової популяції миші Нордмана було затоплено, що може призвести до її зникнення в майбутньому. Знищено до 50% популяції піщаного сліпого землекопа та товстохвостого тушканчика Фальц-Фейна. Дніпровська популяція тритона дунайського теоретично вимерла. Понад 200 особин виду, які занесені до Червоної книги України, опинилися в Чорному морі
Гніздові колонії птахів	У зоні затоплення були знищені важливі місця гніздування водно-болотних і прибережних водоплавних птахів. Це десятки тисяч особин. Саме в заплавах нижнього Дніпра зосереджені найбільші колонії чапель та інших колоніальних птахів регіону
Рослини	Внаслідок підтоплення, швидше за все, загине частина березово-дубових лісів. Крім того, підтоплення призведе до значного підйому рівня ґрунтових вод у всьому південноукраїнському регіоні. Реліктові природні лісові залишки та створені в минулому штучні ліси на дніпровських пісках можуть остаточно зникнути

Рідкісні біотипи	Піщані біотопи, як прибережні, так і континентальні, зі своїми унікальна флора і фауна, найбільше постраждає від затоплення. Крім того, через забруднення води сильно постраждають водні біотопи. Мілкі водойми, які залишаються після спаду рівня води, насправді будуть сумішшю з великою кількістю забруднюючих речовин
Природно-заповідний фонд	Повністю або частково можуть постраждати 48 об'єктів природно-заповідного фонду, з них: один біосферний заповідник; три національні природні парки, один регіональний ландшафтний парк, 16 заказників, три заповідні урочища, 22 пам'ятки природи, два парки-пам'ятки садово-паркового мистецтва
Територія Смарагдової мережі	Втрата природних особливостей цих територій ставить під загрозу виконання Україною своїх зобов'язань щодо збереження цих територій для всієї Європи. Крім того, Козачі острови площею 1000 га віднесені до категорії важливих для збереження птахів
Водно-болотні угіддя міжнародного значення	Дельта Дніпра площею 33 630 га входить до списку охоронюваних територій міжнародного значення згідно з Рамсарською конвенцією. Крім того, опріснення північної частини Чорного моря може негативно вплинути ще на 4 водно-болотні угіддя в регіоні

За даними Міндовкілля, після підриву ГЕС під загрозою знищення опинилися 333 види тварин і рослин різного природоохоронного статусу та 25 типів місць існування. Через цю катастрофу Україна може назавжди втратити деякі екосистеми, зокрема, у національних парках «Нижньодніпровський», «Великий Луг», «Кам'янська Січ» та «Білобережжя Святослава», регіональному ландшафтному парку «Кінбурнська коса», Чорноморському біосферному заповіднику. Опосередковано постраждав національний парк «Олешківські піски» [11].

РОЗДІЛ 2. ДАНІ ДИСТАНЦІЙНОГО ЗОНДУВАННЯ ЯК ОСНОВА ДОСЛІДЖЕННЯ

2.1. Поняття про метод навчання Random Forest для класифікації

У зарубіжній літературі зустрічається доволі багато визначень використаного методу у даному дослідженні, деякі з них наведені нижче.

Випадковий ліс (Random Forest, RF) - це ансамблевий метод навчання для класифікації та регресії. Це важливе досягнення, бо вирішальні дерева, створені в результаті поділу даних, зазвичай недостатньо робастні¹ по відношенню до різних вибірок даних. Тому, взявши дві вибірки, можна отримати два повністю різних дерева класифікації. Це створює серйозні проблеми для перехресної перевірки. Під час ансамблевого навчання створюється декілька вирішальних дерев. Випадковий ліс виправляє схильність вирішальних дерев до перенавчання і, отже, є більше робастною системою класифікації [13].

Випадковий ліс - це сукупність деревних предикторів $h(\mathbf{x}, \Theta_k)$, $k = 1, \dots, K$, де \mathbf{x} представляє спостережуваний вхідний (коваріатний) вектор довжини p з пов'язаним випадковим вектором X , а Θ_k є незалежними та однаково розподіленими випадковими векторами. Передбачається, що спостережувані (тренувальні) дані будуть незалежно отримані від спільного розподілу (X, Y) і включати в себе n $(p + 1)$ – кортежі $(\mathbf{x}_1, y_1), \dots, (\mathbf{x}_n, y_n)$ [14].

Випадковий ліс — це класифікатор, що складається з набору деревоподібних класифікаторів $\{h(\mathbf{x}, \Theta_k), k = 1, \dots\}$, де $\{\Theta_k\}$ є незалежними однаково розподіленими випадковими векторами, і кожне дерево віддає одиничний голос за найпопулярніший клас на вході \mathbf{x} [15].

Примітка. 1. Робастне керування — сукупність методів теорії керування, метою яких є синтез такого контролера або регулятора, який забезпечував би хорошу якість керування, якщо об'єкт керування відрізняється від розрахункового або його математична модель невідома [16].

Метод випадкового лісу являє собою подальше невелике покращення бегінга² дерев рішень, яке полягає в усуненні кореляції між деревами. Як і в випадку із бегінгом, будуються декілька дерев рішень по навчальним бутстреп³-вибіркам. Однак у ході побудови цих дерев перед виконанням кожного розбиття випадковим чином вибирають тільки m із p , які підлягають розгляду предикторів. Відповідне розбиття дозволяється виконувати тільки по одному із цих m предикторів. Для кожного розбиття створюється новий набір із m предикторів, і зазвичай вибирають $m \approx \sqrt{p}$, тобто число предикторів, які розглядаються при кожному розбитті, приблизно рівне квадратному кореню із загального числа предикторів.

Інакше кажучи, під час побудови випадкового лісу під час виконання кожного розбиття алгоритму навіть дозволяється розглядати більшість наявних предикторів. Якщо в даних є один дуже сильний предиктор і кілька інших, які помірно корелюють з відгуком предикторів. Тоді більшість або навіть усі отримані в результаті бегінга дерева у своїй основі використовуватимуть цей сильний предиктор. Отже, всі такі дерева будуть дуже схожі одне на одного. Отримані ж на їх основі прогнози сильно корелюватимуть. На жаль, усереднення великої кількості корелюючих величин не призводить до того ж значного зниження дисперсії, яке відбувається за усереднення некорелюючих величин. Це означає, зокрема, що у такій ситуації бегінг не призведе до суттєвого зниження дисперсії (порівняно з одним деревом) [17, с. 320].

Випадкові ліси вирішують цю проблему, змушуючи алгоритм розглядати перед кожним розбиттям лише певну підмножину предикторів. Отже, в середньому $(p-m)/p$ розбиття повністю проігнорують сильний предиктор, завдяки чому інші предиктори отримують свій шанс.

Примітки. 2. Бегінг є процедурою загального призначення, яка дозволяє знизити дисперсію статистичної моделі. 3. *Статистичний бутстреп* — практичний комп'ютерний метод визначення статистик імовірнісних розподілів, заснований на багаторазовій генерації вибірок методом Монте-Карло на базі наявної вибірки [18].

Можна думати про цю процедуру як усунення кореляції між деревами, завдяки якому усереднення за підсумковими деревами виявляється менш мінливим і, отже, надійнішим.

Основна відмінність між бегінгом та методом RF полягає у виборі розміру підмножини предикторів m . Так, якщо випадковий ліс будується з використанням $m = p$, то вся процедура зводиться до простого бегінга. У порівнянні з бегінгом застосування випадкового лісу з $m \approx \sqrt{p}$ до даних призводить до зниження як помилки на контрольній вибірці, так і помилки на даних, що залишилися (рис. 2.1).

Помилка на контрольних даних (чорна та помаранчева лінії) показана як функція від кількості B^4 навчальних бутстреп-вибірок. Метод випадкового лісу був застосований з використанням $m \approx \sqrt{p}$. Переривчаста лінія показує помилку на контрольній вибірці, отриману внаслідок побудови лише одного дерева. Зелена та блакитна лінії показують помилку на даних, що залишилися (OOB⁵), яка в даному випадку є набагато нижчою.

Вибір малого значення m при побудові випадкового лісу зазвичай буде корисним за наявності великої кількості корелюючих предикторів [17].

Щоб значно підвищити точність класифікації даним методом, було вирощено ансамбль дерев і дозволено їм голосувати за найпопулярніший клас. Для того, щоб виростити ці ансамблі, часто генеруються випадкові вектори, які керують ростом кожного дерева в ансамблі. Раніше згаданим прикладом є бегінг, де для вирощування кожного дерева проводиться випадковий відбір (без заміни) з прикладів з навчальної вибірки. Іншим прикладом є випадковий вибір поділу, коли в кожному вузлі розщеплення вибирається випадковим чином з числа K кращих розщеплень. Брейман генерує нові навчальні набори, рандомізуючи виходи в оригінальній навчальній множині [15].

Примітки. 4. B - кількість дерев, побудованих на основі навчальних бутстреп-вибірок.
5. Решта однієї третини спостережень, які не використовуються для підгонки даного дерева в мішках, називаються спостереженнями поза мішком OOB (out-of-bag).

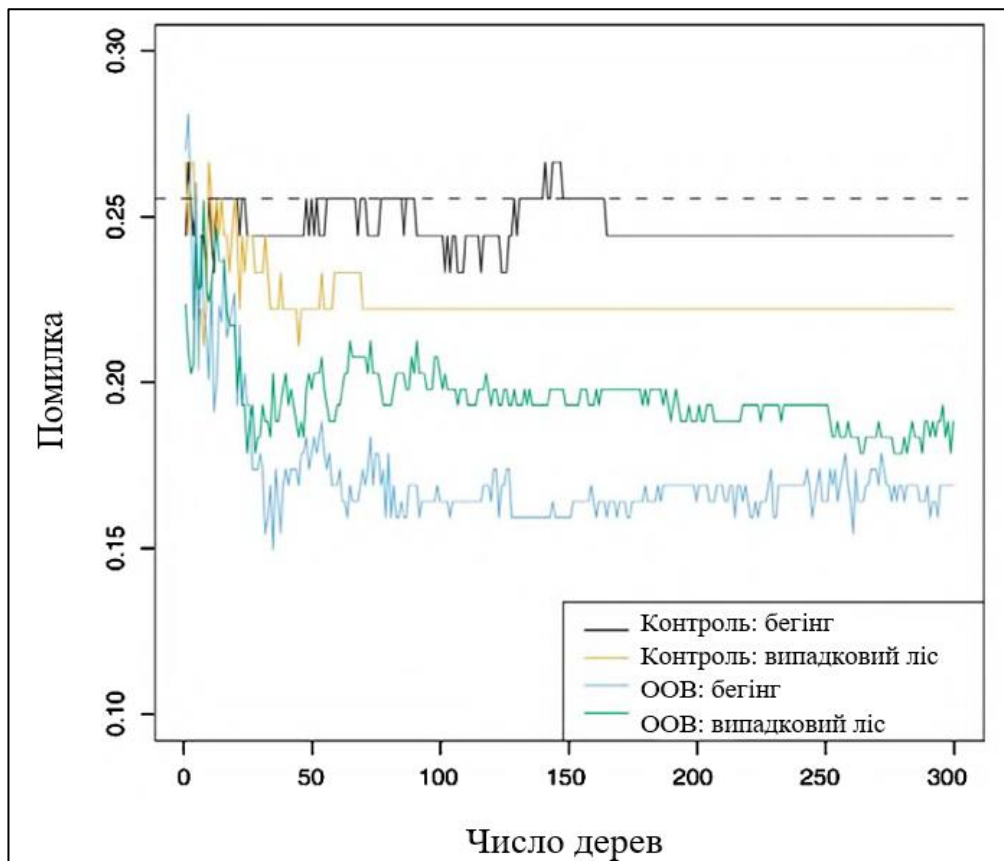


Рис. 2.1. Результати застосування методів бегінга та випадкового лісу до даних

Інший підхід полягає у виборі навчальної множини з випадкового набору ваг на прикладах у навчальній множині. Спільним елементом у всіх процедурах є те, що для k -го дерева генерується випадковий вектор Θ_k , незалежний від минулих випадкових векторів $\Theta_1, \dots, \Theta_{k-1}$, але з тим самим розподілом; і дерево вирощується за допомогою навчальної множини і Θ_k , в результаті чого виходить класифікація $h(\mathbf{x}, \Theta_k)$, де \mathbf{x} — вхідний вектор.

Наприклад, при мішкуванні випадковий вектор Θ генерується як кількість у N коробках, що є результатом N дротиків, кинутих навмання на коробки, де N — кількість прикладів у навчальній множині. При випадковому поділі вибір складається з числа незалежних випадкових цілих чисел від 1 до K .

Характер і розмірність Θ залежить від його використання в деревобудуванні. Після того, як згенеровано велику кількість дерев, вони голосують за найпопулярніший клас. Зазвичай називають ці процедури випадковими лісами.

Дано ансамбль класифікаторів $h_1(\mathbf{x}), h_2(\mathbf{x}), \dots, h_k(\mathbf{x})$, і з навчальною множиною, взятою випадковим чином з розподілу випадкового вектора Y, \mathbf{X} , визначити функцію поля як

$$mg(\mathbf{X}, Y) = av_k I(h_k(\mathbf{X}) = Y) - \max_{j \neq Y} av_k I(h_k(\mathbf{X}) = j), \quad (2.1)$$

де $I(\cdot)$ — індикаторна функція. Маржа вимірює ступінь, в якій середнє число голосів при \mathbf{X}, Y для правого класу перевищує середнє число голосів для будь-якого іншого класу. Чим більша маржа, тим більше довіри до класифікації.

Похибка узагальнення задається формулою

$$PE^* = P_{\mathbf{X}, Y} (mg(\mathbf{X}, Y) < 0), \quad (2.2)$$

де нижні індекси \mathbf{X}, Y вказують на те, що ймовірність знаходиться над простором \mathbf{X}, Y . У випадкових лісах $h_k(\mathbf{X}) = h(\mathbf{X}, \Theta_k)$. Для великого числа дерев з Сильного закону великих чисел і структури дерева впливає, що:

Теорема 1. Зі збільшенням числа дерев майже напевно всі послідовності Θ_1, \dots, PE^* сходяться до

$$P_{\mathbf{X}, Y} (P_{\Theta} (h(\mathbf{X}, \Theta) = Y) - \max_{j \neq Y} P_{\Theta} (h(\mathbf{X}, \Theta) = j) < 0). \quad (2.3)$$

Цей результат пояснює, чому випадкові ліси не перезавантажуються при додаванні більшої кількості дерев, а створюють обмежувальне значення помилки узагальнення. Для випадкових лісів верхня межа може бути виведена для помилки узагальнення в термінах двох параметрів, які є мірою точності окремих класифікаторів і залежності між ними. Взаємодія між ними дає основу для розуміння роботи випадкових лісів.

Визначення 2.1. Функцією запасу для випадкового лісу є

$$mr(\mathbf{X}, Y) = P_{\Theta} (h(\mathbf{X}, \Theta) = Y) - \max_{j \neq Y} P_{\Theta} (h(\mathbf{X}, \Theta) = j) \quad (2.4)$$

та міцністю набору класифікаторів $\{h(\mathbf{x}, \Theta)\}$ є

$$s = E_{X,Y} mr(X, Y). \quad (2.5)$$

Припускаючи $s \geq 0$, нерівність Чебичева дає

$$PE^* \leq var(mr)/s^2. \quad (2.6)$$

Більш наочний вираз для дисперсії mr виводиться в наступному:

$$\hat{j}(X, Y) = \arg \max_{j \neq Y} P_{\Theta} (h(X, \Theta) = j), \quad (2.7)$$

отже

$$\begin{aligned} mr(X, Y) &= P_{\Theta} (h(X, \Theta) = Y) - P_{\Theta} (h(X, \Theta) = \hat{j}(X, Y)) \\ &= E_{\Theta} [I(h(X, \Theta) = Y) - I(h(X, \Theta) = \hat{j}(X, Y))]. \end{aligned} \quad (2.8)$$

Визначення 2.2. Необроблена маржинальна функція має вигляд

$$rmg(\Theta, X, Y) = I(h(X, \Theta) = Y) - I(h(X, \Theta) = \hat{j}(X, Y)). \quad (2.9)$$

Таким чином, $mr(X, Y)$ є математичним сподіванням $rmg(\Theta, X, Y)$ відносно Θ . Для будь-якої функції f тотожність

$$[E_{\Theta} f(\Theta)]^2 = E_{\Theta, \Theta'} f(\Theta) f(\Theta') \quad (2.10)$$

виконується, де Θ, Θ' є незалежними з однаковим розподілом, з чого випливає, що

$$mr(X, Y)^2 = E_{\Theta, \Theta'} rmg(\Theta, X, Y) rmg(\Theta', X, Y). \quad (2.11)$$

Використання (2.11) дає

$$\begin{aligned} var(mr) &= E_{\Theta, \Theta'} \left(cov_{X,Y} rmg(\Theta, X, Y) rmg(\Theta', X, Y) \right) \\ &= E_{\Theta, \Theta'} \left(\rho(\Theta, \Theta') sd(\Theta) sd(\Theta') \right), \end{aligned} \quad (2.12)$$

де $\rho(\Theta, \Theta')$ — кореляція між $rmg(\Theta, X, Y)$ і $rmg(\Theta', X, Y)$ утримує Θ, Θ' фіксований,

$sd(\Theta)$ є стандартним відхиленням $rmg(\Theta, X, Y)$ утримання Θ фіксованим.

Тоді

$$\begin{aligned} var(mr) &= \bar{\rho} (E_{\Theta} sd(\Theta))^2 \\ &\leq \bar{\rho} E_{\Theta} var(\Theta), \end{aligned} \quad (2.13)$$

де $\bar{\rho}$ - середнє значення кореляції, тобто

$$\bar{\rho} = E_{\Theta, \Theta'}(\rho(\Theta, \Theta')sd(\Theta)sd(\Theta'))/E_{\Theta, \Theta'}(sd(\Theta)sd(\Theta')). \quad (2.14)$$

Запишемо

$$\begin{aligned} E_{\Theta} var(\Theta) &\leq E_{\Theta} \left(E_{X, Y} rmg(, X, Y) \right)^2 - s^2 \\ &\leq 1 - s^2. \end{aligned} \quad (2.15)$$

З'єднавши (2.6), (2.13) і (2.15) отримаємо:

Теорема 2. Верхня межа помилки узагальнення задається формулою

$$PE^* \leq \bar{\rho} (1 - s^2)/s^2. \quad (2.16)$$

Хоча межа, швидше за все, буде вільною, вона виконує ту ж сугестивну функцію для випадкових лісів, що і межі типу VC^6 для інших типів класифікаторів. Показано, що двома складовими, які беруть участь у похибці узагальнення для випадкових лісів, є сила окремих класифікаторів у лісі та кореляція між ними з точки зору функцій необробленої межі. Відношення c/s^2 - це кореляція, поділена на квадрат міцності. У розумінні функціонування випадкових лісів це співвідношення буде корисним орієнтиром — чим воно менше, тим краще.

Примітка. 6. VC - це міра ємності (складності, виразної потужності) простору функцій, яких може бути навчено певним алгоритмом статистичної класифікації. Вона визначається як потужність найбільшої множини точок, яку цей алгоритм може розділити [19].

Визначення 2.3. Відношення c/s^2 для випадкового лісу визначається як

$$c/s^2 = \bar{\rho} / s^2. \quad (2.17)$$

У ситуації двох класів є спрощення. Функція *margin* має вигляд

$$mr(\mathbf{X}, Y) = 2 P_{\Theta}(h(\mathbf{X}, \Theta) = Y) - 1. \quad (2.18)$$

Вимога про те, щоб сила була позитивною стає схожою на звично слабку умову навчання $E_{X,Y} P_{\Theta}(h(\mathbf{X}, \Theta) = Y) > .5$. Необроблена функція маржі дорівнює $2I(h(\mathbf{X}, \Theta) = Y) - 1$, а кореляція $\bar{\rho}$ знаходиться між $I(h(\mathbf{X}, \Theta) = Y)$ та $I(h(\mathbf{X}, \Theta') = Y)$. Зокрема, якщо значення для Y приймаються рівними $+1$ і -1 , тоді

$$\bar{\rho} = E_{\Theta, \Theta'} [\rho(h(\cdot, \Theta), h(\cdot, \Theta'))], \quad (2.19)$$

таким чином, $\bar{\rho}$ є кореляцією між двома різними членами лісу, усередненою за Θ, Θ' розподіл.

Для більш ніж двох класів міра сили, визначена в (2.5), залежить як від лісу, так і від окремих дерев, оскільки саме ліс визначає $\hat{j}(\mathbf{X}, Y)$. Можливий і інший підхід. Запишемо

$$\begin{aligned} PE^* &= P_{X,Y} \left(P_{\Theta}(h(\mathbf{X}, \Theta) = Y) - \max_{j \neq Y} P_{\Theta}(h(\mathbf{X}, \Theta) = j) < 0 \right) \\ &\leq \sum_i P_{X,Y} (P_{\Theta}(h(\mathbf{X}, \Theta) = Y) - P_{\Theta}(h(\mathbf{X}, \Theta) = j) < 0). \end{aligned} \quad (2.20)$$

Визначимо

$$s_j = E_{X,Y} (P_{\Theta}(h(\mathbf{X}, \Theta) = Y) - P_{\Theta}(h(\mathbf{X}, \Theta) = j)) \quad (2.21)$$

— сила множини класифікаторів $\{h(\mathbf{x}, \Theta)\}$ відносно класу j . Зазначимо, що це визначення міцності не залежить від лісу. Використовуючи нерівність Чебишева, припускаємо, що всі $s_{j>0}$ призводять до

$$PE^* \leq \sum_i var (P_{\Theta}(h(\mathbf{X}, \Theta) = Y) - P_{\Theta}(h(\mathbf{X}, \Theta) = j)) s_j^2. \quad (2.22)$$

А використовуючи тотожності, подібні до тих, що використовуються при виведенні (2.13), дисперсії в (2.22) можуть бути виражені в термінах середніх кореляцій [15].

Випадкові ліси прагнуть досягти зменшення кореляції шляхом подальшої ін'єкції випадковості. Замість визначення оптимального розбиття даного вузла (складового) дерева шляхом оцінки всіх допустимих розбивок на всі коваріати, як це робиться за допомогою методів одного дерева або мішків, де використовується підмножина коваріат, взятих випадковим чином. Лео Брейман стверджує, що випадкові ліси мають виняткову точність прогнозування, і що ця точність досягається для широкого діапазону налаштувань використовуваного параметра налаштування.

Для регресії випадковим прогнозом лісу є незважене середнє за колекцію:

$$\bar{h}(\mathbf{x}) = (1/K) \sum_{k=1}^K h(\mathbf{x}, \Theta_k). \quad (2.23)$$

Оскільки $k \rightarrow \infty$ закон Великих Чисел забезпечує

$$E_{X,Y} (Y - \bar{h}(X))^2 \rightarrow E_{X,Y} (Y - E_{\Theta} h(X, \Theta))^2 E_{X,Y}. \quad (2.24)$$

Величина праворуч - це похибка прогнозування (або узагальнення) для випадкового лісу, позначена PE_f^* . Збіжність в (2.24) означає, що випадкові ліси не перевміщуються. Тепер визначимо середню похибку прогнозування для окремого дерева $h(X, \Theta)$ як

$$PE_t^* = E_{\Theta} E_{X,Y} (Y - h(X, \Theta))^2. \quad (2.25)$$

Припустимо, що для всіх Θ дерево є незміщеним, тобто, $EY = E_X h(X, \Theta)$. Тоді

$$PE_f^* \leq \bar{\rho} PE_t^*, \quad (2.26)$$

де $\bar{\rho}$ - зважена кореляція між залишками $Y - h(X, \Theta)$ та $Y - h(X, \Theta')$ для незалежних Θ, Θ' .

Нерівність (2.26) вказує на те, що потрібно для точної випадкової регресії лісу:

- низька кореляція між залишками різних деревних членів лісу;
- низька похибка прогнозування для окремих дерев. Крім того, випадковий ліс в очікуванні зменшить індивідуальну помилку дерева, PE_t^* , на коефіцієнт $\bar{\rho}$. Відповідно, введена рандомізація прагне до низької кореляції.

Стратегія, яка використовується для досягнення цих цілей, полягає в наступному:

1. Щоб знизити індивідуальну похибку, потрібно вирощувати дерева на максимальну глибину.
2. Щоб зберегти залишкову кореляцію низькою, варто використовувати рандомізацію за допомогою:
 - а) Вирощення кожного дерева на вибірці початкового завантаження з навчальних даних.
 - б) Використання $m \ll p$ (число коваріат). У кожному вузлі кожного дерева вибирають m коваріат і найкращий розподіл цього вузла на основі цих коваріат [14].

Як приклад, можна розглянути використання методу випадкового лісу до біологічних даних для передбачення типу раку на основі 500 генів, чиї рівні експресії характеризуються найбільшою дисперсією у навчальній вибірці.

Дані містять результати вимірювання рівня експресії 4718 генів у зразках тканини 349 пацієнтів. У цьому наборі даних кожен зразок має якісну мітку з 15 можливими значеннями: нормальна тканина або один із 14 типів раку. Результати представлені на рис. 2.2. Помилка на контрольних даних показана як функція від кількості дерев. Кожна лінія відповідає певному значенню m числа предикторів, доступних для розбиття даних при формуванні кожного внутрішнього вузла дерева. Випадкові ліси ($m < p$) призвели до деякого

поліпшення, порівняно з бегінгом ($m = p$). Частота помилки одного дерева класифікації становить 45,7 %.

400 дерев достатньо для досягнення гарної якості пророцтв і що використання $m \approx \sqrt{p}$ у даному випадку дало незначне покращення помилки на контрольних даних порівняно з бегінгом ($m = p$). Як і у випадку з бегінгом, випадкові ліси не призводять до перенавчання при збільшенні B , і тому на практиці варто вибирати значення, досить високе для стабілізації частоти помилок [17].

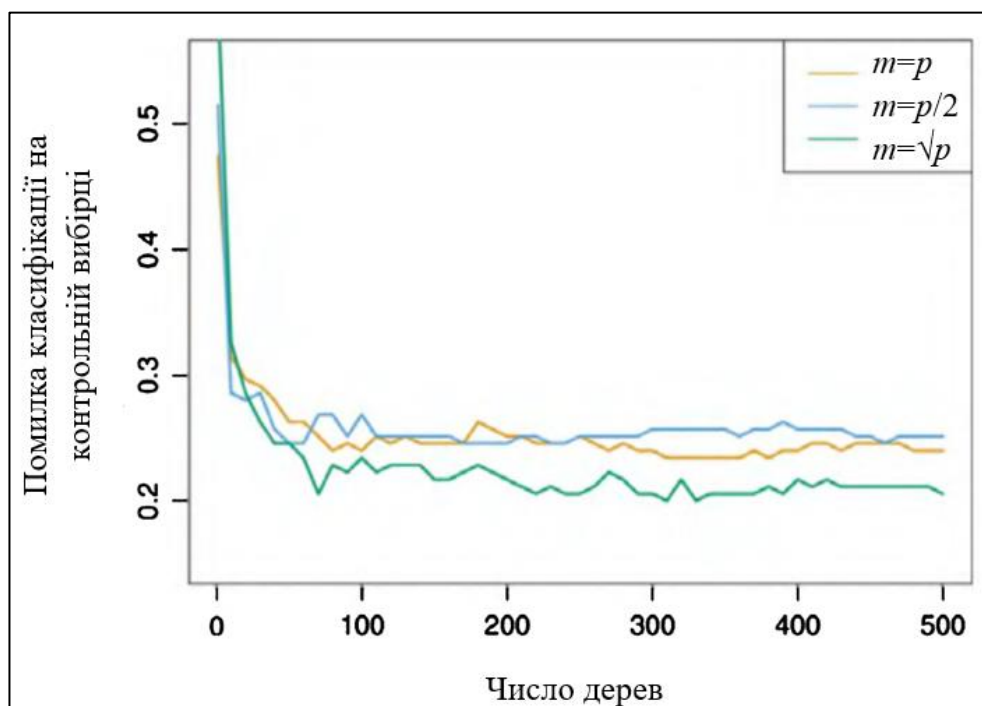


Рис. 2.2. Результати застосування методу випадкового лісу (RF) до даних із 15 класами та 500 предикторами

2.2. Алгоритми процесу класифікації в GEE

Класифікація — це тематична обробка, яка дає можливість проводити автоматизоване розбиття знімків на однорідні за певним критерієм ділянки (класи об'єктів) [20, с. 228].

Класифікація — це процес сортування (розподілу за класами) елементів зображення (пікселів) на кінцеву кількість класів на основі значень їх атрибутів.

Процес класифікації зображень полягає в розподілі всіх пікселів знімка на класи на основі їхньої відбивальної здатності (значення спектральної яскравості) в одній або кількох спектральних зонах, наприклад, для різних типів рослинності або класів використання земель (міські території, пасовища, рілля, водні об'єкти, ліси) [20, с. 380].

У випадку класифікації з навчанням завдання полягає у виявленні на зображенні об'єктів вже відомих типів, що вимагає попередніх знань про досліджувану ділянку земної поверхні. Перший крок процедури передбачає інтерактивний вибір еталонних ділянок на зображенні, які є характерними представниками класів виокремлених об'єктів [20, с. 228].

Для цього можна використати платформу хмарних обчислень Google Earth Engine. Тут можна не тільки завантажувати супутникові дані за вибраний інтервал часу, а й обробляти дані дистанційного зондування та проводити геоінформаційний аналіз. Багато обмежень, пов'язаних із завантаженням, зберіганням і обробкою супутникових даних у різний час, можна легко усунути за допомогою GEE.

Платформа GEE складається з двох основних компонентів: GEE Explorer (для перегляду наборів даних) і GEE Playground, які працюють разом. Доступ до GEE здійснюється через Інтернет, і для роботи з кодом потрібен обліковий запис Google. Каталог даних містить велике сховище загальнодоступних наборів геопросторових даних, включаючи спостереження з різних супутникових і аерофотознімальних систем в оптичних і неоптичних [21].

Як джерело додаткової інформації про землекористування (LULC) при створенні навчальних спектральних сигнатур можна використовувати вегетаційні індекси (VI). При цьому за допомогою них можна визначати не тільки цілеспрямовано окремі типи покриття, а й всі наявні на території дослідження. Їх можна використати ще й як додаткові канали при виконанні класифікації з використанням алгоритмів машинного навчання для підвищення точності класифікації даних Sentinel-2.

Подібність спектрів відбиття поверхонь, які відносяться до штучно створених об'єктів, і поверхонь інших об'єктів з твердою поверхнею, таких як гола земля та пісок, є відомою проблемою, яка виникає під час вилучення LULC за допомогою методів класифікації. Це пов'язано з різноманітністю матеріалів, з яких створюють забудовані території, таких як бетон, асфальт, метал, пластик, скло. Найбільш поширеними є такі VI, як *NDVI*, *NDSI*, *IBI*, *BCI*, *BAEI*, *UI* та *NDWI* [21].

Нормований індекс різниці рослинності (*NDVI*) розраховується за формулою

$$NDVI = \frac{\rho_{NIR} - \rho_{Red}}{\rho_{NIR} + \rho_{Red}} = \frac{Band\ 8 - Band\ 4}{Band\ 8 + Band\ 4}, \quad (2.27)$$

де ρ_{NIR} – значення відбиття довжини хвилі ближнього інфрачервоного діапазону, ρ_{Red} – значення відбиття червоної довжини хвилі.

NDSI розраховується за формулою

$$NDSI = \frac{\rho_{SWIR2} - \rho_{Blue}}{\rho_{SWIR2} + \rho_{Blue}} = \frac{Band\ 12 - Band\ 2}{Band\ 12 + Band\ 2}, \quad (2.28)$$

де ρ_{Blue} – значення відбиття довжини хвилі синього кольору,

ρ_{SWIR2} – значення відбиття довжини хвилі SWIR 2.

IBI розраховується за формулою

$$\begin{aligned} IBI &= \frac{\frac{2 \cdot \rho_{SWIR1}}{\rho_{SWIR1} + \rho_{NIR}} - \left(\frac{\rho_{NIR}}{\rho_{NIR} + \rho_{Red}} - \frac{\rho_{Green}}{\rho_{Green} + \rho_{SWIR1}} \right)}{\frac{2 \cdot \rho_{SWIR1}}{\rho_{SWIR1} + \rho_{NIR}} + \left(\frac{\rho_{NIR}}{\rho_{NIR} + \rho_{Red}} - \frac{\rho_{Green}}{\rho_{Green} + \rho_{SWIR1}} \right)} \\ &= \frac{\frac{2 \cdot Band\ 11}{Band\ 11 + Band\ 8} - \left(\frac{Band\ 8}{Band\ 8 + Band\ 4} - \frac{Band\ 3}{Band\ 3 + Band\ 11} \right)}{\frac{2 \cdot Band\ 11}{Band\ 11 + Band\ 8} + \left(\frac{Band\ 8}{Band\ 8 + Band\ 4} - \frac{Band\ 3}{Band\ 3 + Band\ 11} \right)}, \quad (2.29) \end{aligned}$$

де ρ_{Green} – значення відбиття зеленої довжини хвилі;

ρ_{SWIR1} — значення відбиття довжини хвилі SWIR 1.

BCI розраховується за формулою

$$BCI = \frac{\frac{TC1+TC3}{2} - TC2}{\frac{TC1+TC3}{2} + TC2}, \quad (2.30)$$

де $TC1$, $TC2$ і $TC3$ є компонентами Tasseled Cap для зображень Sentinel-2 MSI.

$BAEI$ розраховується за формулою

$$BAEI = \frac{\rho_{Red} + 0,3}{\rho_{Green} + \rho_{SWIR1}} = \frac{Band\ 4 + 0,3}{Band\ 3 + Band\ 11}. \quad (2.31)$$

UI розраховується за формулою

$$UI = \frac{\rho_{SWIR2} - \rho_{NIR}}{\rho_{SWIR2} + \rho_{NIR}} = \frac{Band\ 12 - Band\ 8}{Band\ 12 + Band\ 8}. \quad (2.32)$$

$NDWI$ розраховується за формулою

$$NDWI = \frac{\rho_{Green} - \rho_{NIR}}{\rho_{Green} + \rho_{NIR}} = \frac{Band\ 3 - Band\ 8}{Band\ 3 + Band\ 8}. \quad (2.33)$$

Процес створення навчальної вибірки (ROI) у GEE полягає в оцифруванні пікселів зображення, які відповідають класу, для якого створюється ROI. Оцифрування можна виконувати за допомогою точок, ліній і багатокутників. Алгоритм перетворює кожен піксель зображення, включений у ROI, на окремий елемент ROI, якщо використовуються лінії або багатокутники. У цьому випадку пікселі області інтересу мають бути максимально однорідними, щоб не допускати помилок у процесі класифікації [21].

Етап навчання полягає фактично в розрахунку й аналізі деякого набору статистичних характеристик розподілу значень пікселів, що складають ці полігони [20, с. 228]. У GEE реалізований метод машинного контрольованого навчання Random Forest (табл. 2.1) [22].

За допомогою комп'ютерної системи пікселі розпізнають і сортують на основі математичного критерію. За теорією розпізнавання образів, простір ознак для класифікації необхідно розділити на замкнені області, кожна з яких містить значення ознак, характерних для одного з класів об'єктів, і призначити кожний піксель знімка тому класу, в ділянку якого потрапив його вектор ознак. Межі, що

розділяють отримані ділянки, називають вирішальними (тими, що поділяють), а самі ділянки — ділянками вирішення [20, с. 381].

Таблиця 2.1

Опис методу Random Forest в GEE

Метод	Аргументи	Значення
ee.Classifier	numberOfTrees (ціле число)	Кількість дерев рішень для створення
.smileRandomForest (Створює порожній класифікатор випадкового лісу і повертає класифіковане зображення)	variablesPerSplit (ціле число, за замовчуванням: null)	Кількість змінних на розбиття. Якщо не вказано, використовується квадратний корінь із кількості змінних
	minLeafPopulation (ціле число, за замовчуванням: 1)	Створює лише вузли, навчальний набір яких містить принаймні таку кількість точок
	bagFraction (число з плаваючою комою, за замовчуванням: 0,5)	Частка входу в «мішок» на дерево
	maxNodest (ціле число, за замовчуванням: null)	Максимальна кількість листових вузлів у кожному дереві. Якщо не вказано, за умовчанням немає обмежень
	seed (ціле число, за замовчуванням: 0)	Засіб рандомізації

Спосіб віднесення пікселів знімка до класів об'єктів визначає так зване вирішальне правило — правило класифікації, реалізацію якого забезпечує відповідний комп'ютерний алгоритм. Правило класифікації формується на основі ознак типових об'єктів, належність яких до певного класу відома. Проте насамперед комп'ютерна система має навчитися розпізнавати образи на основі таких відомостей, і в процесі навчання потрібно знайти критерій розпізнавання образів [20, с. 381].

При контрольованій класифікації передбачаються правила переходу від показників спектрального відбиття до положення в системі класів об'єктів, перевіряються вони на «навчальній» (еталонній) ділянці. Ці правила

виробляються на основі ознак навчальної вибірки, введених фахівцем, а потім автоматично застосовуються і на іншій частині знімка. Використовуючи контрольовану класифікацію важливо знати кількість і тип класів, що виокремлюються, і лише потім створювати навчальні вибірки, що відповідають відомостям знімка [20, с. 381].

Методи контрольованої класифікації враховують апріорну (вже відому) інформацію про:

- типи об'єктів та ймовірності їхнього представлення даними знімка;
- еталонні значення спектральних характеристик цих об'єктів (еталонні значення їхнього спектрального образу), від якості яких залежить точність і надійність класифікації [20, с. 382].

Часто виникає потреба в тематичному коректуванні результатів класифікації, особливо автоматичному, виконуваному фактично на основі інформаційних характеристик об'єктів. Для цього використовують набір процедур, який називають операціями посткласифікаційної обробки: злиття класів, розділення класів, усунення дрібних помилкових об'єктів, згладжування меж об'єктів тощо [20, с. 229].

Весь процес можна виконати у будь-якій ГІС програмі, зокрема в ArcGIS Pro. ArcGIS Pro, потужний однокористувацький настільний додаток ГІС, являє собою багатофункціональне програмне забезпечення, розроблене з урахуванням удосконалень та ідей спільноти користувачів. Це програмне забезпечення підтримує візуалізацію даних, розширений аналіз та авторитетне ведення даних у 2D, 3D та 4D.

З його допомогою дані перетворюються на карти та корисну інформацію, забезпечується можливість комбінування даних з різних джерел для створення гарних та професійних карт, а також публікації проектів та інформації в межах організації, для користувачів інтернету або мобільних додатків. Інтерактивна графіка та параметри аналізу можуть редагуватися на льоту. Інтерактивні інструменти допомагають створювати об'єкти аналітики простими клацаннями

миші у сцені або під час використання вхідних шарів джерела, забезпечуючи візуальний зворотний зв'язок у реальному часі [23].

Для великих класифікованих зображень краще використовувати такі інструменти:

- «Region Group» (Група регіонів). Для кожної комірки у вихідних даних записується ідентичність підключеної області, до якої ця комірка належить [24].
- «Set Null» (Встановити Null). «Set Null» встановлює значення NoData для визначених клітинок на основі визначених критеріїв [25].
- «Nibble» (Відсікання). Замінює комірки растру, що відповідає масці, значеннями найближчих сусідів [26].

Важливим етапом у процесі класифікації є оцінка точності отриманих зображень [20, с. 229]. Щоб провести її в GEE можна скласти матрицю плутанини, згідно з якою розраховувати найбільш часто використовувані описові статистики для оцінки класифікації: загальна точність (OA), точність виробника (PA), точність користувача (UA) і Коефіцієнт Каппа (Каппа) [21].

Загальна точність обчислюється шляхом підсумовування великої діагоналі матриці помилок (тобто правильно відображених одиниць вибірки) і ділення на загальну кількість одиниць вибірки в матриці, вона означає точність усієї карти. Важливо, щоб матриця помилок була повідомлена, щоб можна було обчислити інші показники точності.

Однак, крім того, потрібно знати точність окремого класу тематичної карти. Точність виробника та користувача обчислюється шляхом ділення значення великої діагоналі (узгодження) для цього класу на загальну кількість зразків у цьому класі карти. Оцінюючи точність окремого класу карти, важливо враховувати як точність виробника, так і користувача.

Загалом, точність виробника та користувача також можна пояснити за допомогою наступних рівнянь. Починають з n вибірок, які розподіляються на k^2 клітинок, де кожна вибірка призначається одному з k класів тематичної карти на

карті (зазвичай рядки) i , незалежно, одному з тих же k класів тематичної карти в наборі довідкових даних (зазвичай колонки). Отже, нехай n_{ij} позначає кількість зразків, відображених у класі тематичної карти i ($i = 1, 2, \dots, k$) у карті та класі тематичної карти j ($j = 1, 2, \dots, k$) встановлені у довідкових даних.

Отримаємо

$$n_{i+} = \sum_{j=1}^k n_{ij}, \quad (2.34)$$

які є кількістю зразків, віднесених до класу тематичної карти « i » на карті, і

$$n_{+j} = \sum_{i=1}^k n_{ij}, \quad (2.35)$$

які є кількістю зразків, віднесених до класу тематичної карти « j » у наборі довідкових даних. Тоді загальну точність між картою та довідковими даними можна обчислити наступним чином:

$$\text{Загальна точність} = \frac{\sum_{i=1}^k n_{ii}}{n}. \quad (2.36)$$

Точність виробника обчислюється за формулою

$$\text{Точність виробника } j = \frac{n_{ii}}{n_{+j}} \quad (2.37)$$

і точність користувача може бути обчислена

$$\text{Точність користувача } i = \frac{n_{ii}}{n_{i+}}. \quad (2.38)$$

Другий метод, який часто використовувався для оцінки точності, називається «Каппа (Карра)». Результати виконання Каппа-аналізу є статистикою КНАТ, яка може бути використана як інший показник узгодженості або точності. Статистика КНАТ обчислюється як

$$\hat{K} = \frac{\sum_{i=1}^r x_{ii} - \sum_{i=1}^r (x_{i+} \cdot x_{+i})}{N^2 - \sum_{i=1}^r (x_{i+} \cdot x_{+i})}, \quad (2.39)$$

де r — кількість рядків у матриці,

x_{ii} — кількість спостережень у рядку i та стовпці i ,
 x_{i+} та x_{+i} — граничні підсумки рядка i та стовпця i ,
 N — загальна кількість спостережень.

Каппа опосередковано включає помилки, використовуючи суми підсумків рядків і стовпців у обчисленні статистики. Таким чином, кожна міра представляє різну інформацію і повинна бути оцінена відповідним чином.

Лендіс і Кох запропонували певний діапазон значень статистики Каппа для представлення різних рівнів узгодженості. Значення більше 80 % означає високу згоду, діапазон між 60 % і 80 % означає помірну згоду, а значення нижче 40 % – слабку згоду. Однак основний внесок Каппа-аналізу полягає не як міра точності, а як техніка, яка використовується для визначення того, чи одна матриця помилок є статистично значущою від іншої [27, с. 595-597].

2.3. Дані як основа дослідження

У даній роботі було використано космічні знімки компанії Copernicus, інформація про них наведена нижче.

Місія Sentinel-1 є Європейською радіолокаційною обсерваторією для спільної ініціативи Європейської комісії (ЕК) та Європейського космічного агентства (ESA) Copernicus. Copernicus – це європейська ініціатива щодо впровадження інформаційних послуг, що стосуються навколишнього середовища та безпеки. Він базується на даних спостережень, отриманих із супутників спостереження за Землею, і на наземній інформації.

Місія Sentinel-1 складається з двох сонячно-синхронних супутників на полярній орбіті, які мають одну площину орбіти з різницею фаз орбіти 180°, працюють вдень і вночі, виконуючи радіолокаційні зображення з синтетичною апертурою С-діапазону, що дозволяє їм отримувати зображення незалежно від погоди. Він працює в чотирьох ексклюзивних режимах зйомки з різною роздільною здатністю (до 5 м) і покриттям (до 400 км) та забезпечує подвійну поляризацію, дуже короткий час повторного перегляду та швидку доставку

продукту. Для кожного спостереження доступні точні вимірювання положення та орієнтації космічного корабля.

Sentinel-1 є першою з п'яти місій, розроблених ESA для ініціативи Copernicus. Його домен вимірювання охоплює рельєф ландшафту, багатоцільові зображення (земля), багатоцільові зображення (океан), вітри на поверхні океану, рельєф/течії океану, висоту та спектр океанських хвиль, морський льодовий покрив, край і товщину, сніговий покрив, край і глибина, вологість ґрунту та рослинність.

Супутник оснащений двома крилами сонячної батареї, здатними виробляти 5900 Вт (наприкінці терміну служби), які зберігатимуться в модульній батареї, ємність батареї становить 324 А·год.

Sentinel-1 знаходиться на близькополярній сонячно-синхронній орбіті з 12-денним циклом і 175 обертами за цикл для одного супутника. I Sentinel-1A, і Sentinel-1B мають однакову площину орбіти з різницею орбітальних фаз у 180°. Коли обидва супутники працюють, цикл повторення становить шість днів (рис. 2.3).

Один супутник Sentinel-1 потенційно здатний скласти карту всього світу в інтерферометричному режимі Wide swath раз на 12 днів із 175 обертами за цикл (рис. 2.4). Оскільки відстань між траєкторіями орбіти змінюється залежно від широти, швидкість повторного відвідування (зростання/спадання) значно вища на вищих широтах, ніж на екваторі: менше ніж 1 день в Арктиці та забезпечує покриття Європи та Канади, основні судноплавні маршрути за 1-3 днів і через 3 дні на екваторі. Радарні дані доставляються до служб Copernicus протягом години після отримання.

Прилад забезпечує можливість отримання зображень у будь-якій погоді, вдень і вночі, щоб отримувати дані вимірювань із високою та середньою роздільною здатністю для спостереження за сушею, прибережними зонами та льодом.

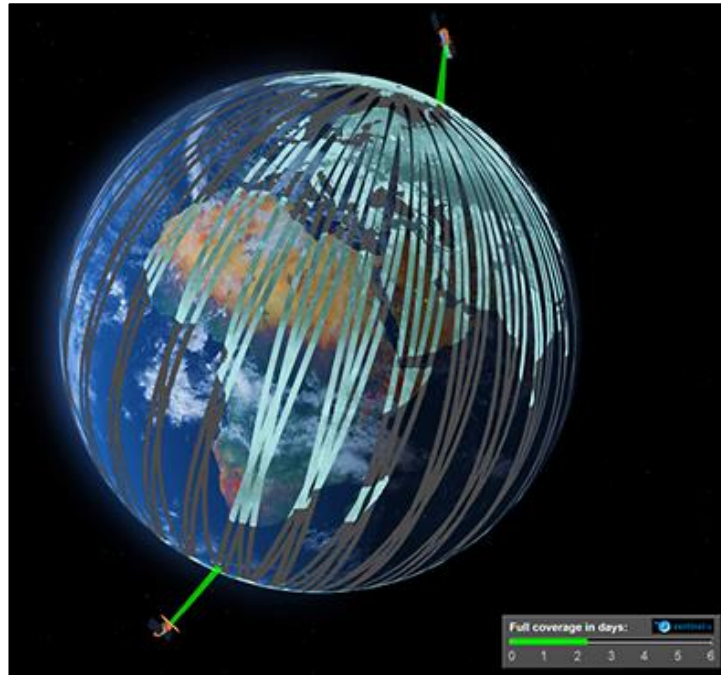


Рис. 2.3. Сузір'я Sentinel-1. Автори: ESA/ATG medialab

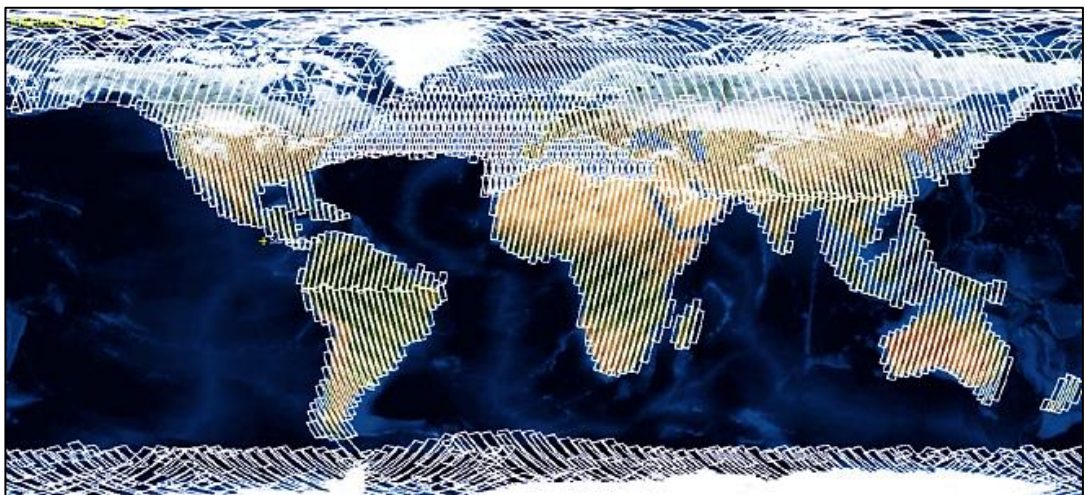


Рис. 2.4. Потенційне покриття світу Sentinel-1 IW

Інструмент C-SAR — це активна фазована антенна решітка, що забезпечує швидке сканування по куту кута (щоб охопити великий діапазон кутів падіння та підтримувати роботу ScanSAR) і по азимуту (щоб дозволити використовувати техніку TOPS для досягнення необхідної якості зображення) з подвійною ємністю (табл. 2.2). Він має внутрішню схему калібрування, коли сигнали передачі направляються в приймач, щоб забезпечити моніторинг амплітуди/фази для забезпечення високої радіометричної стабільності.

Основні параметри приладу Sentinel-1 C-SAR

Характеристики	Значення
Центральна частота	5,405 ГГц (довжина хвилі ~5,55 см)
Пропускна здатність	0-100 МГц
Поляризація	HH+HV, VV+VH, VV, HH
Розмір антени	12,3 м × 0,821 м
Пікова потужність РЧ (сума всіх TRM, при TRM о/р)	4,368 кВт, -4,075 кВт
Ширина імпульсу	5-100 мкс
Робочий цикл передачі	Макс. 12 %, SM 8,5 %, IW 9 %, EW 5 %, WV 0,8 %
Коефіцієнт шуму приймача на вході модуля	3,2 дБ
Частота дискретизації АЦП	300 МГц
Загальна маса приладу (включаючи антену)	945 кг
Ефективна швидкість низхідного каналу CSAR	Два канали по 260 Мбіт/с

На відміну від пасивних оптичних датчиків, яким потрібне сонячне освітлення, активний прилад SAR передає власний мікрохвильовий сигнал, щоб освітлювати поверхню Землі під кутом. SAR активно передає мікрохвильові сигнали на Землю та отримує частину переданої енергії у вигляді зворотного розсіювання від землі.

Прилад SAR забезпечує вимірювання зворотного розсіювання радіолокації під впливом структури рельєфу та нерівності поверхні. Як правило, чим більше нерівностей або структури на землі, тим більше зворотне розсіювання. Плоскі поверхні відбивають сигнал, що призводить до темних елементів. Подібним чином складніші за структурою цілі, такі як ліси, виглядатимуть яскравіше, оскільки взаємодія сигналу з листям, гілками та стовбурами призведе до того, що більша частка сигналу передається назад на датчик.

Sentinel-1 — система SAR із подвійною поляризацією із збереженням фази. Він може передавати сигнал у горизонтальній (H) або вертикальній (V) поляризації, а потім приймати як у H, так і V поляризації.

Сигнал SAR містить інформацію про амплітуду та фазу. Амплітуда – це сила відгуку радара, а фаза – частка одного повного циклу синусоїдальної хвилі (одна довжина хвилі SAR). Фаза PCA-зображення визначається насамперед відстанню між супутниковою антеною та наземними цілями [28].

Sentinel-2 — це європейська широкозонна місія з високою роздільною здатністю багатоспектральних зображень. Повна специфікація місії супутників-близнюків, що летять на одній орбіті, але з фазою 180° , розроблена таким чином, щоб забезпечити високу частоту повторного відвідування екватора 5 днів.

Sentinel-2 несе корисне навантаження оптичного приладу, який знімає 13 спектральних смуг: чотири смуги 10 м, шість смуг 20 м і три смуги з просторовою роздільною здатністю 60 м. Ширина орбітальної смуги становить 290 км (рис. 2.5).

Ці супутники використовуються для підтримки різноманітних послуг і додатків, зокрема управління землею, сільське господарство, лісове господарство, боротьба зі стихійними лихами, операції з надання гуманітарної допомоги, картографування ризиків і проблеми безпеки.

Місія Sentinel-2 складається з двох ідентичних супутників Sentinel-2A і Sentinel-2B, які були запуснені за допомогою європейської ракети-носія VEGA.

Два супутники Sentinel-2 працюють одночасно, фазовані під кутом 180° один до одного, на сонячно-синхронній орбіті на середній висоті 786 км. Положення кожного супутника Sentinel-2 на його орбіті вимірюється двочастотним приймачем глобальної навігаційної супутникової системи (GNSS).

MSI працює пасивно, збираючи сонячне світло, відбите від Землі. Прилад отримує нові дані, коли супутник рухається по своїй орбітальній траєкторії. Вхідний світловий промінь розділяється через світлорозділювач і фокусується на двох окремих вузлах фокальної площини всередині приладу; один для видимого та ближнього інфрачервоного (VNIR) діапазонів, а інший для короткохвильового інфрачервоного (SWIR) діапазонів.

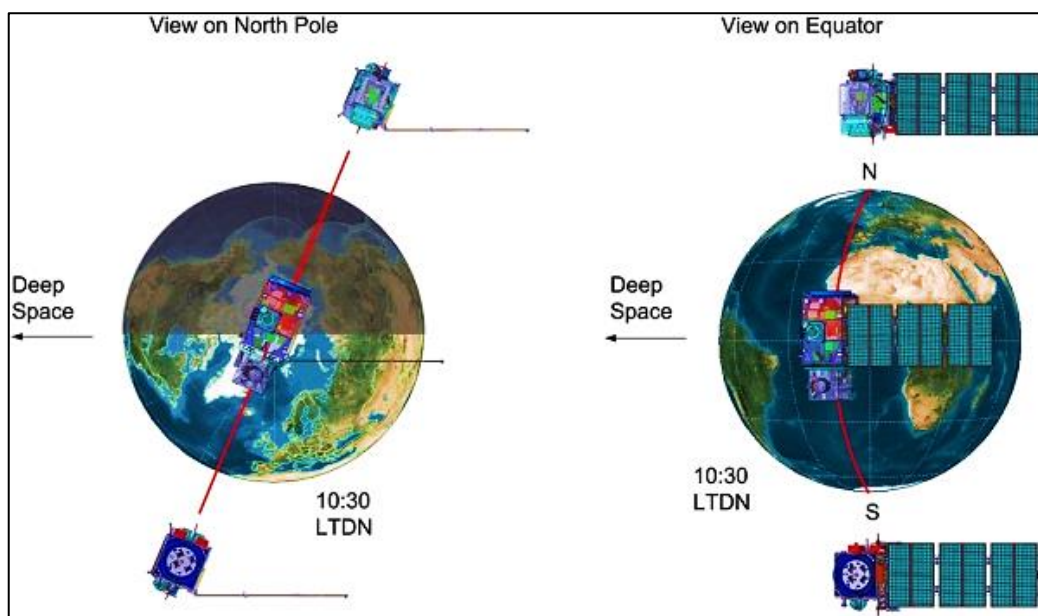


Рис. 2.5. Орбітальна конфігурація подвійного супутника Sentinel-2. Автор:
Astrium GmbH

Орбіта місії Sentinel-2 є сонячно-синхронною. Сонячно-синхронні орбіти використовуються для забезпечення незмінного кута сонячного світла на земну поверхню. Крім невеликих сезонних коливань, прив'язка орбіти супутників до кута сонця мінімізує потенційний вплив тіней і рівнів освітлення на землю. Це забезпечує послідовність у часі та має вирішальне значення для оцінки даних часових рядів.

Sentinel-2A і Sentinel-2B займають одну орбіту, розділену 180 градусами. Середня орбітальна висота становить 786 км. Нахил орбіти становить $98,62^\circ$, а середній місцевий сонячний час (MLST) у низхідному вузлі становить 10:30 (ранку). Це значення MLST було обрано як компроміс між відповідним рівнем сонячного освітлення та мінімізацією потенційної хмарності. Значення близьке до локального часу естакади Landsat і майже ідентичне SPOT-5, що дозволяє інтегрувати дані Sentinel-2 з існуючими та історичними місіями та сприяє збору довгострокових часових рядів даних.

MultiSpectral Instrument (мультиспектральний інструмент, MSI) в супутниках використовує концепцію push-broom. Датчик push-broom працює, збираючи ряди даних зображення по всій орбітальній смужці та використовує рух

космічного корабля вперед уздовж траєкторії орбіти, щоб створити нові рядки для отримання даних.

Спектральна роздільна здатність визначається як міра здатності приладу розрізняти особливості в електромагнітному спектрі. 13 спектральних діапазонів Sentinel-2 варіюються від видимого (VNIR) та ближнього інфрачервоного (NIR) до короткохвильового інфрачервоного (SWIR):

- 4 x 10-метрові діапазони: три класичні діапазони RGB (синій (~493 нм), зелений (560 нм) і червоний (~665 нм)) і ближній інфрачервоний діапазон (~833 нм) (рис. 2.7);
- 6 x 20-метрових смуг: 4 вузькі смуги в спектральному домені червоного краю рослинності VNIR (~704 нм, ~740 нм, ~783 нм і ~865 нм) і 2 ширші смуги SWIR (~1610 нм і ~2190 нм) для таких застосувань, як сніг/лід/виявлення хмар або оцінка стресу від вологості рослинності (рис. 2.8);
- 3 x 60-метрові смуги в основному зосереджені на скринінгу хмар і атмосферній корекції (~443 нм для аерозолів і ~945 нм для водяної пари) і виявленні перистих (~1374 нм) (рис. 2.9).

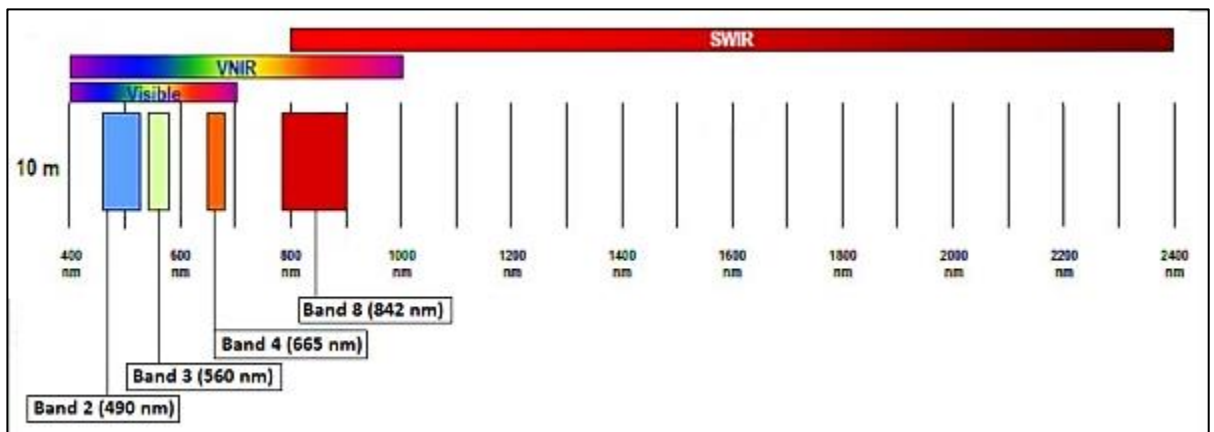


Рис. 2.7. Смуги просторової роздільної здатності Sentinel-2 10 м: B2, B3, B4 і B8

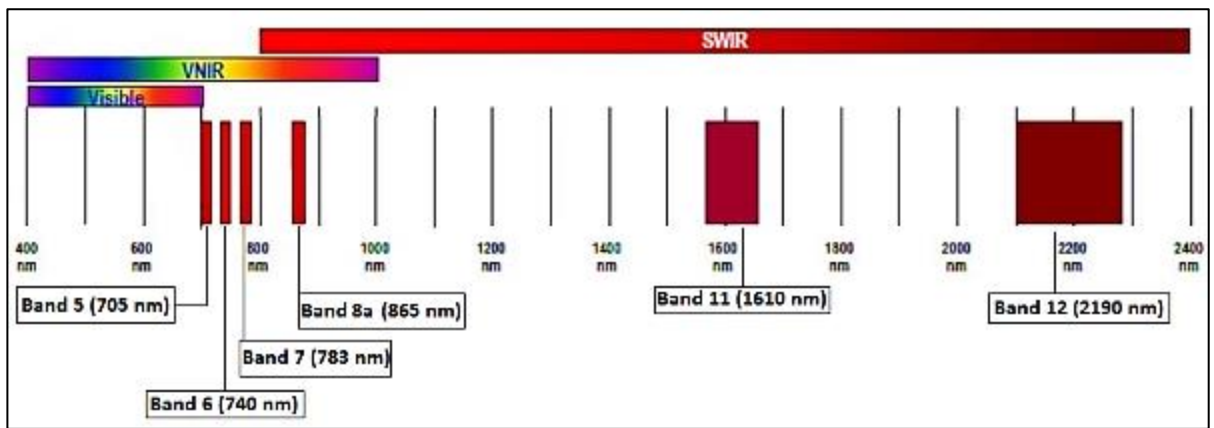


Рис. 2.8. Смуги просторової роздільної здатності Sentinel-2 20 м: B5, B6, B7, B8a, B11 і B12

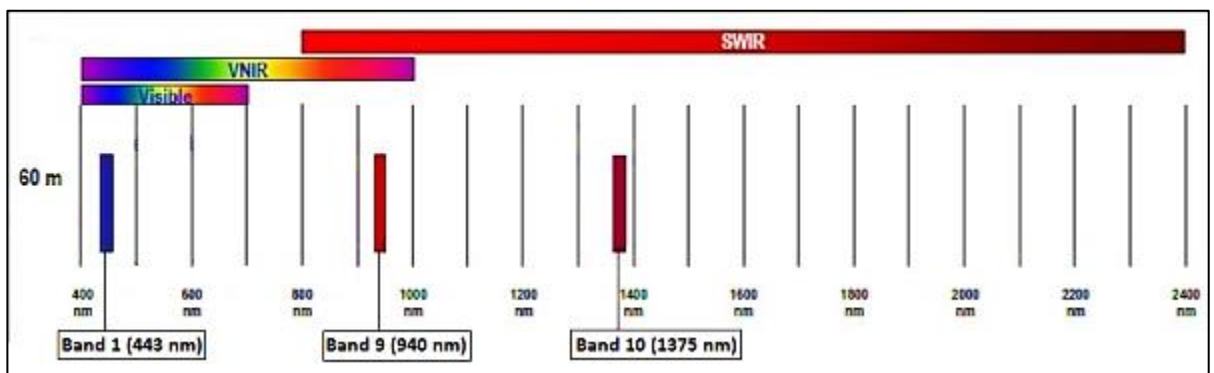


Рис. 2.9. Смуги просторової роздільної здатності Sentinel-2 60 м: B1, B9 і B10

Радіометрична роздільна здатність — це здатність приладу розрізняти різницю в інтенсивності світла або коефіцієнті відбиття. Чим більше радіометрична роздільна здатність, тим точнішим буде отримане зображення.

Еквівалентна довжина хвилі відома з похибкою нижче 1 нм. Таким же чином помилка вимірювання абсолютного калібрування впливає на точність фізичного значення.

Місія Sentinel-2 систематично збирає дані про сушу та прибережні райони в смузі широт, що простягається від 56° південної (острів Горнос, мис Горн, Південна Америка) до 82,8° північної (над Гренландією):

- всі прибережні води до 20 км від берега;
- всі острови площею понад 100 км²;
- всі острови ЄС;

- Середземне море;
- всі закриті моря (наприклад, Каспійське море).

Крім того, сценарій спостереження Sentinel-2 включає спостереження за запитамі держав-членів або служб Copernicus (наприклад, Антарктида, Баффінова затока) [29].

Додатковим джерелом даних стали топографічні карти досліджуваної території за 1921, 1923, 1941, 1988, 1990 рр.

РОЗДІЛ 3. АНАЛІЗ ОТРИМАНИХ РЕЗУЛЬТАТІВ

3.1. Класифікація знімків методом RF в GEE

Перед класифікацією було створене безхмарне мозаїчне зображення Sentinel-2 у середніх значеннях спектрального відбиття (рис. 3.1) за 20.05. – 05.06.2023 рр. (перед підливом ГЕС, далі комбінація 1), 01.08. – 31.08.2023 рр. (період, коли зникла вода або залишилась місцями в незначній кількості, далі комбінація 2) та 21.05-25.05.2024 рр. (період майже через рік, далі комбінація 3) для постраждалої території Херсонської області, а саме міста Херсон, Олешки, Нова Каховка та інших прибережних населених пунктів річки Дніпра (рис. 3.2). Аргументи «start», «end» являють собою періоди, для яких виконується класифікація, наприклад, 20.05 - «start», 05.06 – «end».

Для маскуванню хмар було завантажено колекцію Cloud Score+, яка є процесором оцінки якості (QA) для оптичних супутникових зображень середньої та високої роздільної здатності. Набір даних Cloud Score+ S2_HARMONIZED оперативно створюється з узгодженої колекції Sentinel-2, а результати Cloud Score+ можна використовувати для ідентифікації відносно чітких пікселів і ефективного видалення хмар і тіней від хмар L1C (верхня атмосфера) або L2A.

```
var csPlus = ee.ImageCollection('GOOGLE/CLOUD_SCORE_PLUS/V1/S2_HARMONIZED');
var QA_BAND = 'cs_cdf';
// The threshold for masking; values between 0.50 and 0.65 generally work well.
// Higher values will remove thin clouds, haze & cirrus shadows.
var CLEAR_THRESHOLD = 0.60;

function maskS2clouds(image) {
  var qa = image.select('QA60');

  // Bits 10 and 11 are clouds and cirrus, respectively.
  var cloudBitMask = 1 << 10;
  var cirrusBitMask = 1 << 11;

  // Both flags should be set to zero, indicating clear conditions.
  var mask = qa.bitwiseAnd(cloudBitMask).eq(0)
  | | .and(qa.bitwiseAnd(cirrusBitMask).eq(0));

  return image.updateMask(mask).divide(10000);
}
```

Рис. 3.1. Процес маскуванню хмар (код із GEE)

```

var dataset = ee.ImageCollection('COPERNICUS/S2_SR_HARMONIZED')
    .filterDate(start, end)
    .linkCollection(csPlus, [QA_BAND])
    .map(function(img) {
    return img.updateMask(img.select(QA_BAND)
    .gte(CLEAR_THRESHOLD));
    })
    // .filter(ee.Filter.lt('CLOUDY_PIXEL_PERCENTAGE', 5))
    .filterBounds(boundary)
    .map(maskS2clouds);

var composite = dataset.median();

```

Рис. 3.2. Завантаження мультиспектральних даних та створення мозаїки (код із GEE)

Для підвищення точності класифікації були завантажені колекції радарних даних Sentinel-1 (рис. 3.3). Дані містять різні поляризації, тому знімки було відфільтровано за ними та оброблено, а саме усунено шуми та згладжено за смугами VV, VH. Після чого з них були створені мозаїки за ті ж періоди, що і мультиспектральні (рис. 3.4).

```

var S1_VV = ee.ImageCollection('COPERNICUS/S1_GRD')
    .filterBounds(boundary)
    .filterDate(start, end)
    .filter(ee.Filter
    .listContains('transmitterReceiverPolarisation', 'VV'))
    .map(function(image) {
    var edge = image.lt(-30.0);
    var maskedImage = image.mask().and(edge.not());
    return image.updateMask(maskedImage);
    });

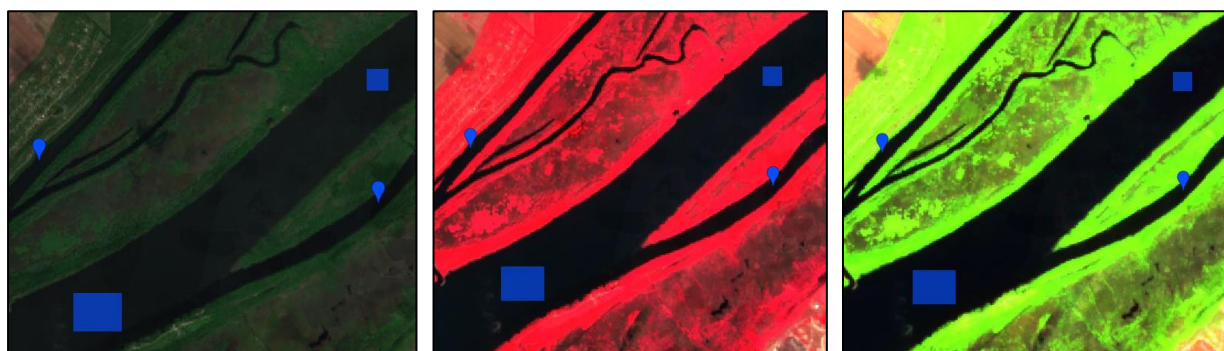
```

```

var S1_VH = ee.ImageCollection('COPERNICUS/S1_GRD')
    .filterBounds(boundary)
    .filterDate(start, end)
    .filter(ee.Filter
    .listContains('transmitterReceiverPolarisation', 'VH'))
    .map(function(image) {
    var edge = image.lt(-30.0);
    var maskedImage = image.mask().and(edge.not());
    return image.updateMask(maskedImage);
    });

```

Рис. 3.3. Завантаження радарних даних (код із GEE)

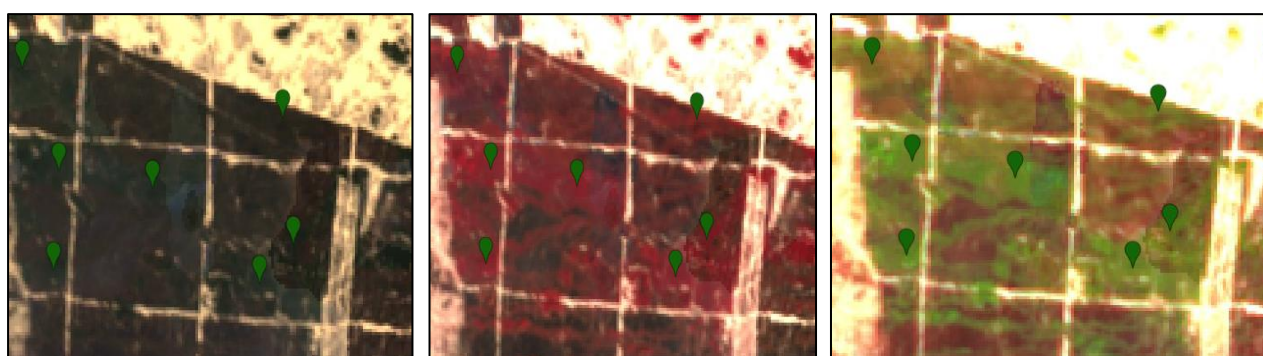


а

б

в

Рис. 3.5. Приклад навчальних вибірок для води (сині точки та полігони) в різних комбінаціях каналів: а - RGB = 4:3:2, б - RGB = 8:4:3, в - RGB = 11:8:4

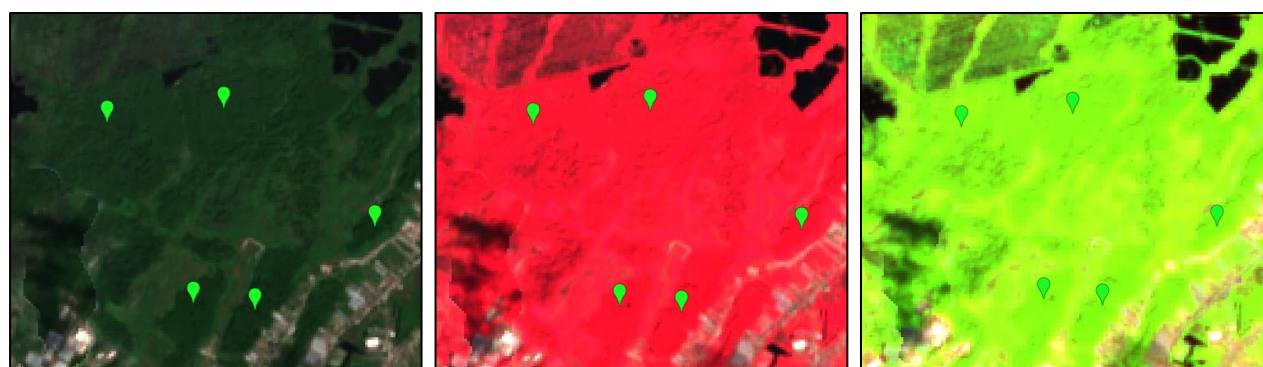


а

б

в

Рис. 3.6. Приклад навчальних вибірок для лісу (темно-зелені точки) в різних комбінаціях каналів: а - RGB = 4:3:2, б - RGB = 8:4:3, в - RGB = 11:8:4



а

б

в

Рис. 3.7. Приклад навчальних вибірок для рослинності (яскраво-зелені точки) в різних комбінаціях каналів: а - RGB = 4:3:2, б - RGB = 8:4:3, в - RGB = 11:8:4

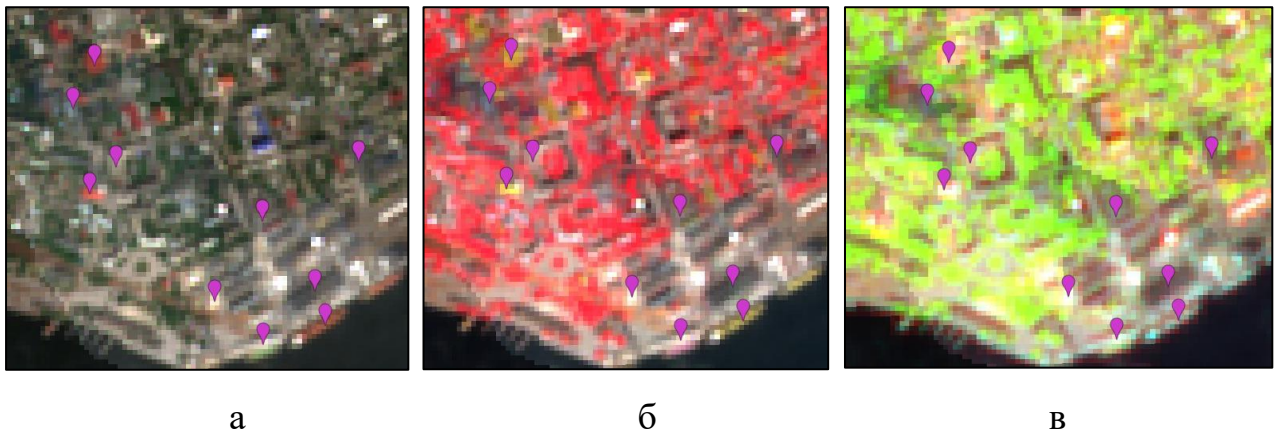


Рис. 3.8. Приклад навчальних вибірок для населених пунктів (фіолетові точки) в різних комбінаціях каналів: а - $RGB = 4:3:2$, б - $RGB = 8:4:3$, в - $RGB = 11:8:4$

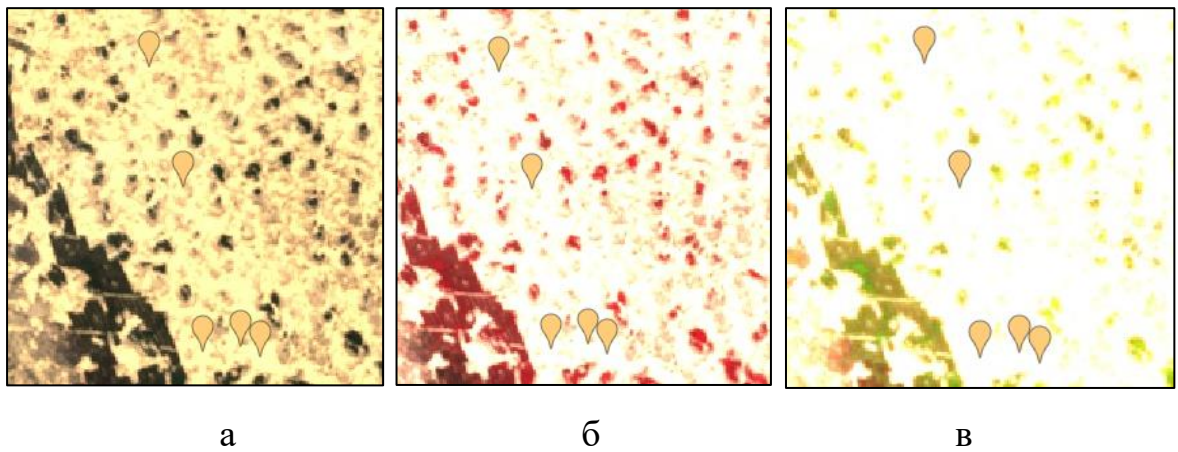


Рис. 3.9. Приклад навчальних вибірок для піску (бежеві точки) в різних комбінаціях каналів: а - $RGB = 4:3:2$, б - $RGB = 8:4:3$, в - $RGB = 11:8:4$



Рис. 3.10. Приклад навчальних вибірок для відкритих ґрунтів (жовті точки) в різних комбінаціях каналів: а - $RGB = 4:3:2$, б - $RGB = 8:4:3$, в - $RGB = 11:8:4$

Таким чином, створено всього 378, 476 та 477 навчальних вибірок для комбінацій 1, 2 та 3 відповідно, включаючи полігональні об'єкти. Кількість дерев рішень, які потрібно створити, є обов'язковим входним параметром для радіочастотної роботи над GEE, їх було використано 100.

Для підвищення точності перед процесом класифікації було об'єднано мультиспектральні та радарні дані (рис. 3.12), створено одну навчальну вибірку «sample», яка містить в собі всі раніше створені області інтересу (рис. 3.13).

Коли були підготовлені космічні знімки, навчальні вибірки та встановлені необхідні параметри, був запущений процес класифікації за допомогою алгоритмів машинного навчання RF. На рис. 3.14 наведений код тренування для цього методу, а на рис. 3.15 запущена сама класифікація.

```
var S1_S2 = ee.Image.cat(composite.select('B1', 'B2', 'B3', 'B4', 'B5', 'B7',  
                                         'B8', 'B8A', 'B9', 'B11', 'B12'),  
                        VV, VH, VV_VH, ui_index, ndsi, ibi, bci, baei);  
  
var bands_Sentinel_1_2 = ['B1', 'B2', 'B3', 'B4', 'B5', 'B7', 'B8', 'B8A', 'B9', 'B11', 'B12',  
                          'VV_Filtered', 'VH_Filtered', 'VVVH_Filtered',  
                          'ui_index', 'ndsi', 'ibi', 'bci', 'baei'];
```

```
var training_S1_S2 = S1_S2.select(bands_Sentinel_1_2).sampleRegions({  
  collection: sample,  
  properties: ['class'],  
  scale: 10 });
```

Рис. 3.12. Об'єднання оптичних та радарних даних (код із GEE)

```
var sample = water.merge(forest).merge(vegetation)  
            .merge(urban).merge(sands).merge(bare_soils);
```

Рис. 3.13. Об'єднання навчальних вибірок (код із GEE)

```
var classifier_S1_S2_RF = ee.Classifier.smileRandomForest(100).train({  
  features: training_S1_S2,  
  classProperty: 'class',  
  inputProperties: bands_Sentinel_1_2  
});
```

Рис. 3.14. Тренування класифікатора RF (код із GEE)

```
var classified_S1_2_RF = S1_S2.select(bands_Sentinel_1_2).classify(classifier_S1_S2_RF);
```

Рис. 3.15. Класифікація об'єднаних каналів (код із GEE)

Після завершення були отримані класифіковані зображення. Територія дослідження досить велика, щоб перевіряти кожен класифікований піксель самостійно, тому було вирішено скласти матрицю плутанини для оцінки точності отриманих зображень. Для цього попередньо створено 284, 280 та 312 вибірок, включно з полігональними, для перевірки комбінації 1, 2 та 3. Таким же чином вибірки було об'єднано в одну – «valNames» (рис. 3.16). Матриці було складено та пораховано статистики безпосередньо в GEE (рис. 3.17). На основі цього з отриманих результатів були сформовані табл. 3.1, 3.2, 3.3.

```
var valNames = vWater.merge(vForest).merge(vVegetation)
               .merge(vUrban).merge(vSands).merge(vBare_soils);
```

Рис. 3.16. Об'єднання вибірок для перевірки точності (код із GEE)

```
var validation = classified_S1_2_RF.sampleRegions({
  collection: valNames,
  properties: ['class'],
  scale: 10,
});
print(validation);

var testAccuracy = validation.errorMatrix('class', 'classification');
//Print the error matrix to the console
print('Validation error matrix: ', testAccuracy);
//Print the overall accuracy to the console
print('Validation overall accuracy: ', testAccuracy.accuracy());

var confusionMatrix = testAccuracy;
print("Constructed confusion matrix", confusionMatrix);
// Calculate overall accuracy.
print("Overall accuracy", confusionMatrix.accuracy());
// Calculate consumer's accuracy
print("Consumer's accuracy", confusionMatrix.consumersAccuracy());
// Calculate producer's accuracy
print("Producer's accuracy", confusionMatrix.producersAccuracy());
// Calculate kappa statistic
print('Kappa statistic', confusionMatrix.kappa());
```

Рис. 3.17. Оцінка точності класифікації (код із GEE)

Матриця плутанини показує, наскільки точно модель класифікує різні типи землекористування, порівнюючи передбачені та фактичні значення. Правильно визначені пікселі розташовуються по діагоналі, наприклад, для класу водні об'єкти правильно класифіковані всі 751 піксель. Неправильно визначеними пікселями є, наприклад, 3 пікселі класу ліс помилково віднесено до класу відкритих ґрунтів.

Таблиця 3.1

Матриця плутанини для RF LULC (період 20.05–05.06.2023)

	Водні об'єкти	Ліс	Рослин- ність	Населені пункти	Пісок	Відкриті грунти	Всього	Точність користу- вача (UA, %)
Водні об'єкти	751	0	0	0	0	0	751	100
Ліс	0	53	0	0	0	1	54	98,15
Рослинність	0	0	51	0	0	0	51	100
Населені пункти	0	0	0	63	0	0	63	100
Пісок	0	0	0	1	51	0	52	98,08
Відкриті грунти	0	3	0	1	0	48	52	92,31
Всього	751	56	51	65	51	49	1023	
Точність виробника (РА, %)	100	94,64	100	96,92	100	97,96		
Загальна точність (ОА, %)	99,41							
Каппа	0,99							

Таблиця 3.2

Матриця плутанини для RF LULC (період 01.08–31.08.2023)

	Водні об'єкти	Ліс	Рослин- ність	Населені пункти	Пісок	Відкриті грунти	Всього	Точність користу- вача (UA, %)
Водні об'єкти	2818	0	0	0	0	0	2818	100
Ліс	0	50	0	0	0	2	52	96,15
Рослинність	0	0	52	0	0	0	52	100
Населені пункти	1	1	2	49	2	3	58	84,48
Пісок	0	0	0	0	51	0	51	100
Відкриті грунти	0	0	0	0	0	57	57	100
Всього	2819	51	54	49	53	59	3088	
Точність виробника (РА, %)	99,96	98,08	96,3	100	96,23	91,94		
Загальна точність (ОА, %)	99,64							
Каппа	0,97							

Матриця плутанини для RF LULC (21.05-25.05.2024)

	Водні об'єкти	Ліс	Рослин- ність	Населені пункти	Пісок	Відкриті грунти	Всього	Точність користу- вача (UA, %)
Водні об'єкти	1315	0	0	0	0	0	1315	100
Ліс	0	49	0	0	0	1	50	98
Рослинність	0	0	52	0	0	0	52	100
Населені пункти	0	0	4	63	0	0	67	94,03
Пісок	0	0	0	0	64	0	64	100
Відкриті грунти	0	0	0	0	0	52	52	100
Всього	1315	49	56	63	64	53	1600	
Точність виробника (РА, %)	100	100	92,86	100	100	98,11		
Загальна точність (ОА, %)	99,63							
Каппа	0,99							

Звідки, загальна точність класифікацій 99,41 %, 99,64 % та 99,63 %, оцінка Каппа – 0,99, 0,97 та 0,99 відповідно. Найбільше помилок в усіх випадках виникало при ідентифікації населених пунктів, іноді і рослинність, і пісок відображались як міста/села. В усіх випадках правильно відображались водні об'єкти, рослинність, іноді траплялись похибки із лісом та піском. Але збільшивши кількість навчальних вибірок, точність стала більшою.

У першій комбінації всі класи, крім піску та відкритих ґрунтів, не мають похибок взагалі. У другій – ситуація схожа, але вже населені пункти і пісок мають помилково ідентифіковані пікселі. Третя комбінація має схожість із другою, але містить похибки в класі пісок. Лише у другій комбінації з класом пісок 100% точність визначення моделлю.

Найменші точності виробника в класах ліс, відкриті ґрунти та рослинність, а користувача – населені пункти, відкриті ґрунти. Відкриті ґрунти найважче розпізнавались і користувачем, і виробником.

Отже, отримані результати можна вважати точними та використовувати в подальшому. Загалом модель має високу точність як для виробника, так і для користувача, а також загальну точність. Значення Каппа також свідчить про високу узгодженість моделі.

Впевнившись в достовірності отриманих класифікованих зображень, їх було завантажено для посткласифікаційної обробки в ArcGIS Pro. Для обробки було використано інструменти «Region Group», «Set Null» та «Nibble» з модуля «Spatial Analyst», які усунули шуми та неправильно класифіковані поодинокі пікселі. У тому ж програмному середовищі були створені компоновки отриманих карт LULC (рис. 3.16, 3.17, 3.18).

Досить помітним є зміна рослинності на відкриті ґрунти (див. рис. 3.16, 3.17), це пов'язано з різними часовими відрізками отриманих комічних даних для комбінацій 1 та 2. У травні все здебільшого покрите рослинністю через активний період росту культур, а в серпні вже з більшості полів зібрано урожаї і переорано до наступного сезону.

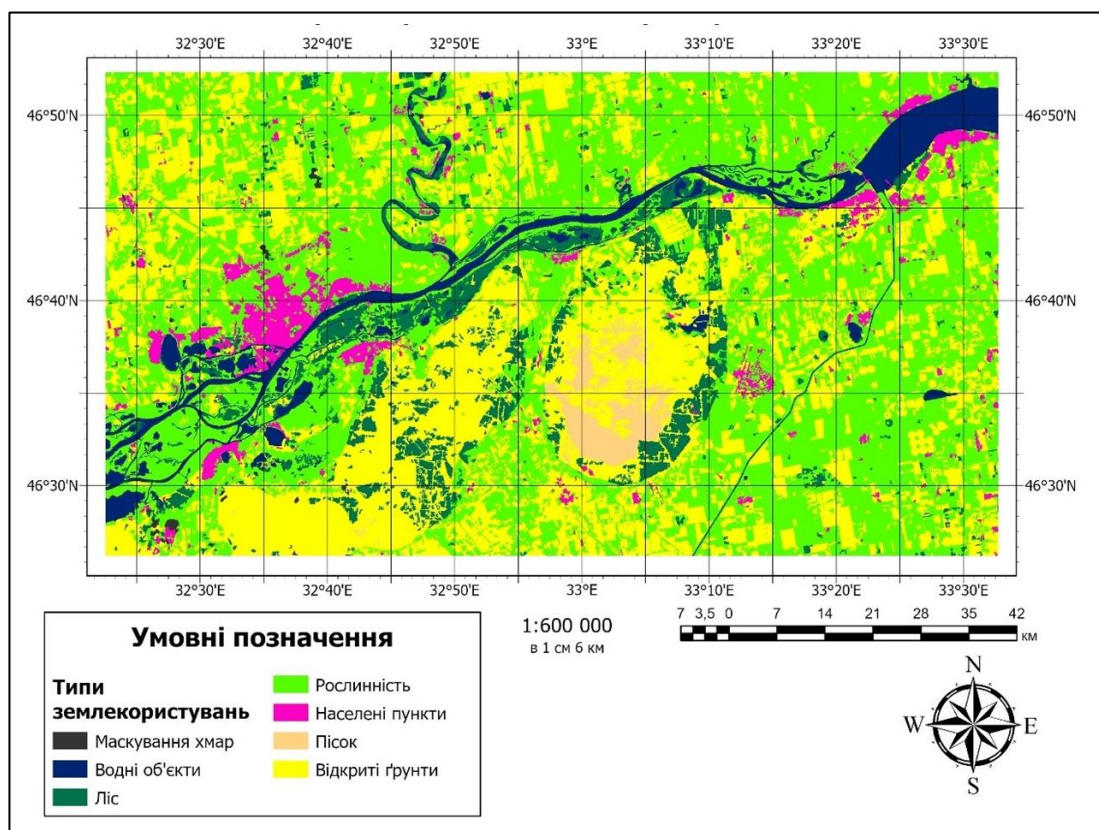


Рис. 3.16. LULC для 20.05-05.06.2023 pp.

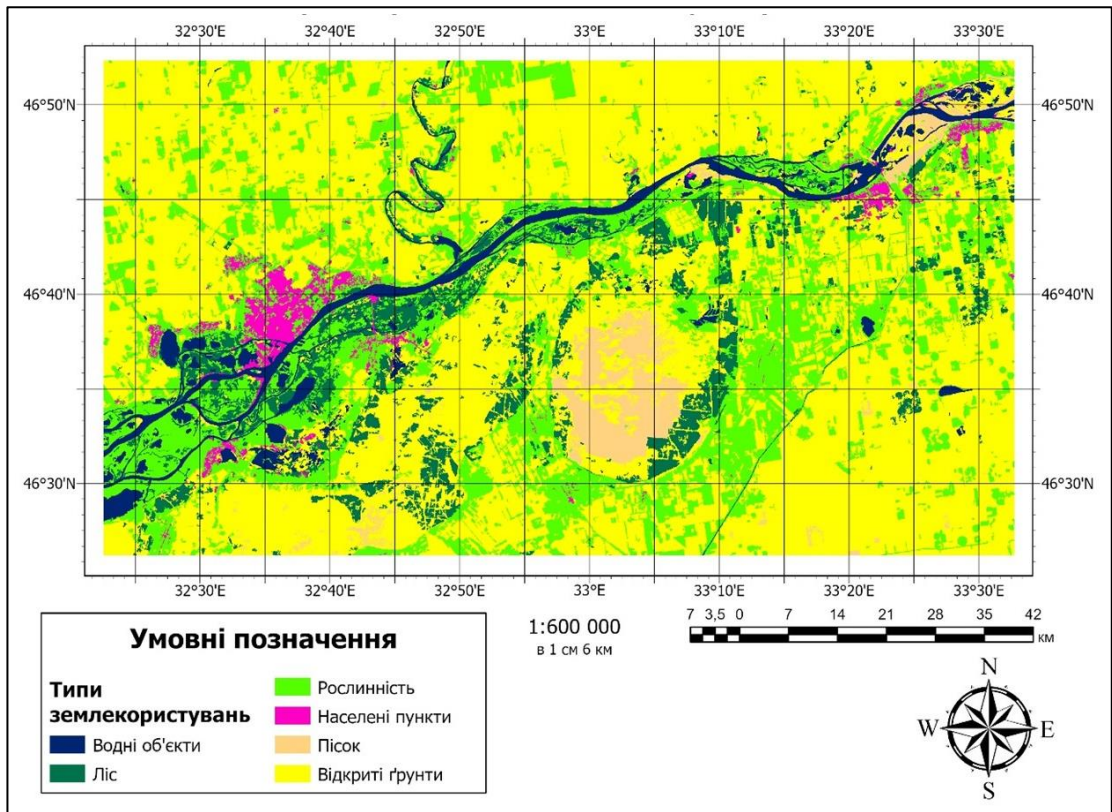


Рис. 3.17. LULC для 01.08-31.08.2023 рр.

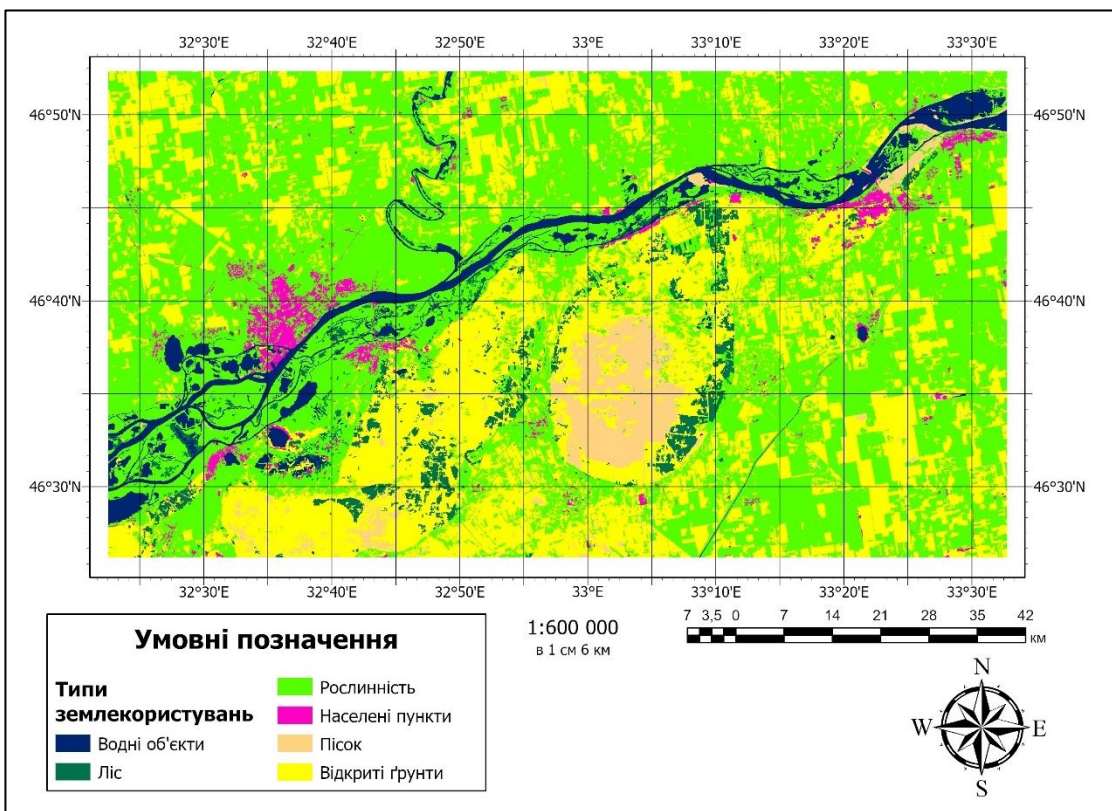


Рис. 3.18. LULC для 21.05-25.05. 2024 рр.

Схожа ситуація, але навпаки, між комбінаціями 2 і 3 (див. рис. 3.17, 3.18). Це зумовлено, знову ж так, різними періодами отриманих даних та активним ростом культур. Помітні також зміни в кількості лісів, їх значно менше, ніж за попередній рік, що говорить про велику вирубку хвойних дерев в даному регіоні.

Також добре видно масштаби зменшення площі води в районі Каховського водосховища (у правому верхньому куті). Система вод Дніпра повернулась майже до рівня до 1956 року. Однак майже через рік рівень води піднявся. Якщо взяти до уваги карти ще по побудови Каховської ГЕС Дніпро виглядало так, як на рис. 3.19. А вже в 1989 році мало вигляд такий, як на рис. 3.20.

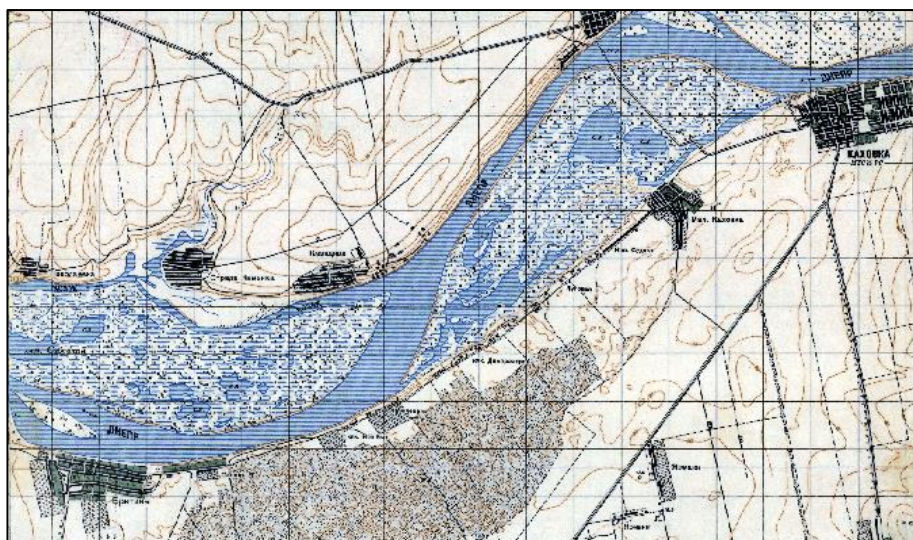


Рис. 3.19. Топографічна карта міста Каховка за 1941 р.



Рис. 3.20. Топографічна карта міста Каховка за 1989 р.

Для більш кращого візуального ефекту було взято класифіковане зображення комбінації 2 і топографічну карту за 1941 р. і за допомогою функції «Layer Blend» – «Exclusion» відображено різницю між двома растровими зображеннями (рис. 3.21). Різними кольорами відображені типи землекористувань, а вигляд топографічної карти зберігся для тих ділянок, які змінились, а самі контури карти – це межа змін після 1941 року.

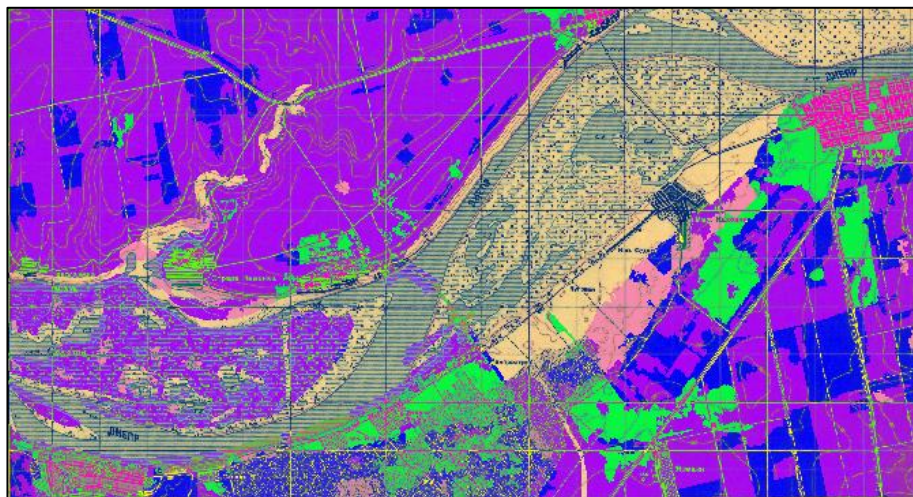


Рис. 3.21. Різниця між топографічною картою (1941 р.) та класифікованим зображенням (20.05-05.06.2023)

Із рисунку видно, що острівна частина Дніпра, до місця розташування майбутньої ГЕС, вся пішла під воду, при цьому затопило і ділянки, які раніше використовувались. Інші острови (лівий ніжній край) не так сильно пішли під воду. Крім цього збільшився об'єм рівня води річки з часом, яка впадала в Дніпро. Отже, найбільші зміни сталися на території до лінії побудови дамби, греблі та станції в місті Нова Каховка.

А якщо взяти комбінацію 3 і ту саму топографічну карту за 1941 р., можна побачити наскільки річка відновилася (рис. 3.22). Місцями рівень води залишився таким, як і в серпні минулого року, а місцями (верхня частина карти) острівна частина Дніпра майже повністю така, якою і була раніше. Рівень води опустився до колишнього рівня, це видно по лівій прибережній зоні.

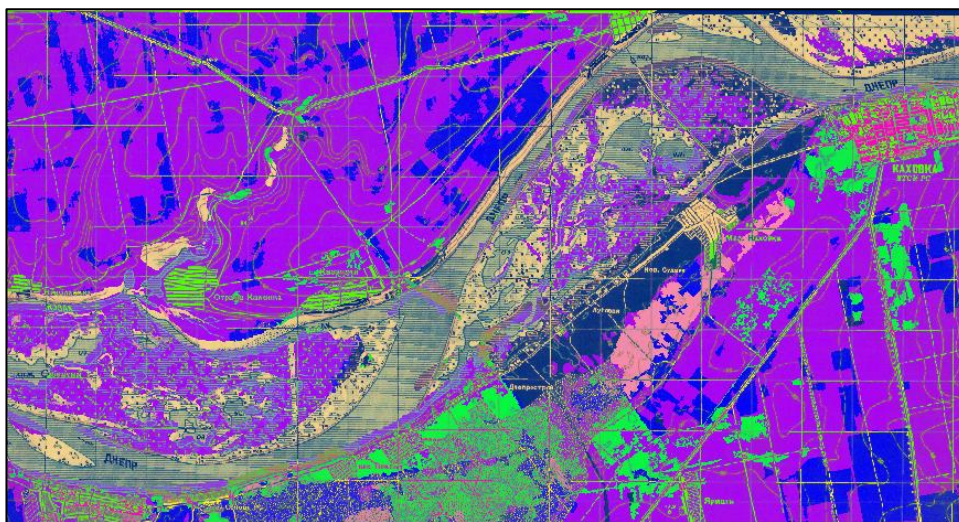


Рис. 3.22. Різниця між топографічною картою (1941 р.) та класифікованим зображенням (21.05-25.05.2024)

3.2. Виявлення затоплених територій за допомогою радарних даних

З урахуванням об'ємів води, які після підриву, затопили прибережні території, важливим було знайти та виявити ці затоплені території. Для цього були використані радарні дані Sentinel-1. Знімки взято за період 01.04 – 04.06.2023 рр. для знімку «before» (мозаїка даних, де немає затоплених територій, рис. 3.23) та окремі знімки на кожну дату окремо від 9 червня по 27 липня 2023 року (рис. 3.24). Використовувалась лише одна поляризація VV.

```
var collection = ee.ImageCollection('COPERNICUS/S1_GRD')
  .filterBounds(boundary)
  .filter(ee.Filter.listContains('transmitterReceiverPolarisation', 'W'))
  .select('VV')

//filter the before and after flood imagery
var before = collection.filterDate('2023-04-01', '2023-06-04').mosaic()
```

Рис. 3.23. Завантаження колекції (код із GEE)

```
var after_0906 = ee.Image('COPERNICUS/S1_GRD/S1A_IW_GRDH_1SDV_20230609')
var after_1406 = ee.Image('COPERNICUS/S1_GRD/S1A_IW_GRDH_1SDV_20230614')
var after_2106 = ee.Image('COPERNICUS/S1_GRD/S1A_IW_GRDH_1SDV_20230621')
var after_2606 = ee.Image('COPERNICUS/S1_GRD/S1A_IW_GRDH_1SDV_20230626')
var after_0307 = ee.Image('COPERNICUS/S1_GRD/S1A_IW_GRDH_1SDV_20230703')
var after_0807 = ee.Image('COPERNICUS/S1_GRD/S1A_IW_GRDH_1SDV_20230708')
var after_1507 = ee.Image('COPERNICUS/S1_GRD/S1A_IW_GRDH_1SDV_20230715')
var after_2007 = ee.Image('COPERNICUS/S1_GRD/S1A_IW_GRDH_1SDV_20230720')
var after_2707 = ee.Image('COPERNICUS/S1_GRD/S1A_IW_GRDH_1SDV_20230727')
```

Рис. 3.24. Завантаження радарних знімків (код із GEE)

Знімки було обрізано відповідно до досліджуваної території та оброблено, а саме усунено шум (рис. 3.25). Тепер потрібно знайти різницю між мозаїкою «before» та знімками вже за конкретні дати. Вже потім переглядаючи карти «difference» підібрати відповідні коефіцієнти, які відповідатимуть саме за воду (рис. 3.26) – це і будуть повені, які можна використати як маску, щоб залишились лише ці ділянки (рис. 3.27).

```
var before_s = before_clip.focal_median(60, 'circle', 'meters')
var after_s1 = after_0906.focal_median(60, 'circle', 'meters')
```

Рис. 3.25. Застосування фільтру згладжування (код із GEE)

```
var difference = after_s1.subtract(before_s)
```

```
var flood_extent = difference.lt(-3)
```

Рис. 3.26. Знаходження різниці між знімками до і після (код із GEE)

```
var flood = flood_extent.updateMask(flood_extent)
```

Рис. 3.27. Маскування затоплених територій (код із GEE)

Отримані ділянки було завантажено для постобробки в ArcGIS Pro, щоб, знову ж таки, прибрати шуми і зайві ділянки, які не стосуються води. Адже при такому алгоритмі пошуку повеней різниця знаходиться для всього і можуть бути помилково за них взяті звичайні поля. Звичайно, щоб не видалити і потрібні ділянки, проводилось порівняння радарних даних із отриманими.

У результаті отримано таку карту – рис. 3.28, решта карт наведена в додатках А, Б, В, Г, Д, Е, Ж, И. Вони дозволяють наочно зрозуміти масштаби і приблизно оцінити площі територій, які певний час знаходились з водою. Однак для більш точного результату отримані растрові дані було перетворено у полігональні об'єкти. За допомогою функції «Calculate Geometry» в таблиці атрибутів пораховані площі. Графік розподілу можна побачити на рис. 3.29.

За 09.06 (3 день після підриву) найбільше територій зазнали повені, потім обсяг в 2,2 рази зменшився, потім в 2,4 в порівнянні з 16.06. У 4 рази зменшились

площі на 03.07 в порівнянні з 26.06. Потім поступово до позначки 5,66 км² (27.07) і до кінця місяця практично вода була лише на острівній частині Дніпра.

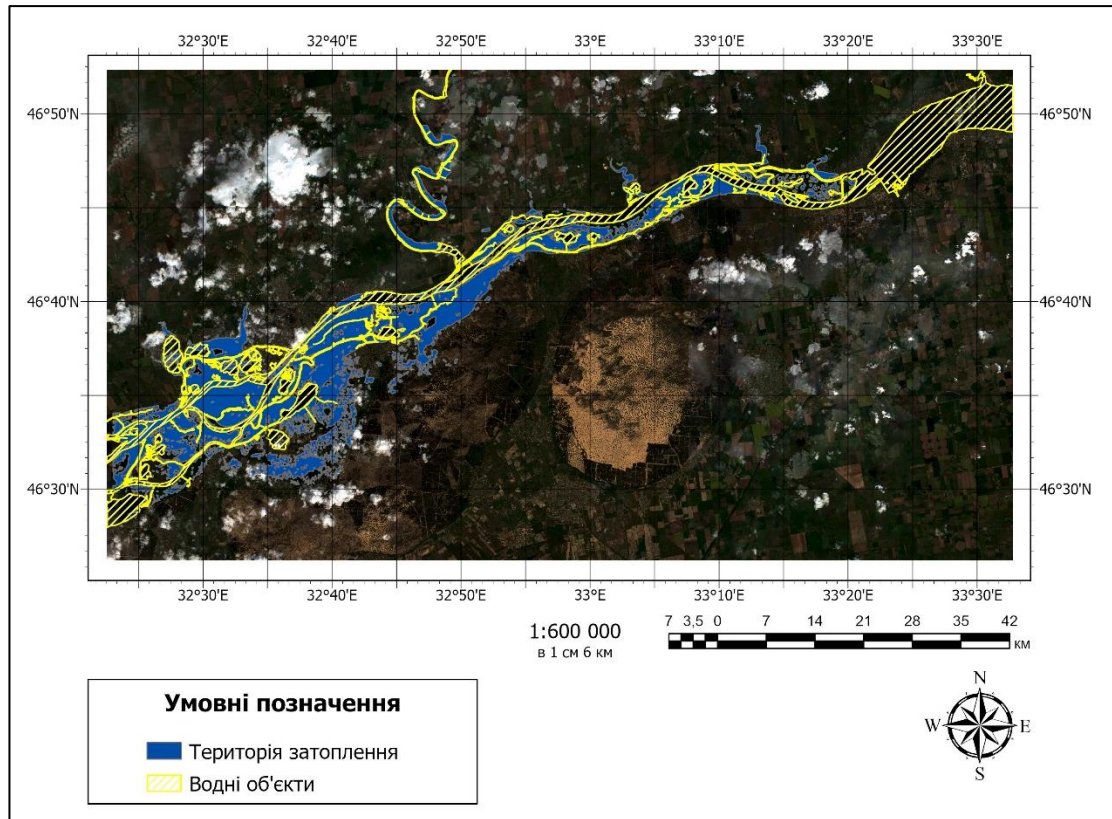


Рис. 3.28. Затоплені території після підриву за 09.06.2023



Рис. 3.29. Лінійний графік площі затоплених територій за датами

У липні спостерігались деякі збільшення площ, зайнятих водою, це може бути зумовлено збільшенням кількості опадів цього місяця.

3.3. Аналіз змін типів землекористувань

На отриманих картах LULC вже візуально видно певні зміни в землекористуванні, особливо в районі Каховської ГЕС, де від попереднього водосховища залишилась майже первісна система вод Дніпра. Після підриву дамби води водосховища зменшилися настільки, що залишились тільки пісок і ґрунт. Якщо порівняти це із даними топографічної карти за 1941 р. рівень води в цьому районі був значно більшим.

Щоб більш точно оцінити кількісну зміну між картами комбінацій 1 і 2 (рис. 3.30, 3.31) в програмному середовищі ArcGIS Pro за допомогою інструментів було пораховано площі змін землекористувань. Для цього класифіковані зображення були перетворені в полігональні об'єкти, потім за допомогою інструментів «Dissolve», «Intersect» знайдені зміни між класами. Потім функцією «Calculate Geometry» пораховані їх площі.

Найбільше змінилися рослинності на відкриті ґрунти – 1313,84 км², відкриті ґрунти на рослинність – 321,32 км². Найбільші площі, які залишились без змін: відкриті ґрунти – 1131,18 км², рослинність – 587,09 км². Досить велику площу і надалі займають піски – 74,47 км², вони практично залишились у попередньому вигляді, лише незначні 0,2 км² стали відкритими ґрунтами.

Відбуваються значні перетворення лісових площ в рослинність - 38,58 км² і відкриті ґрунти – 75,06 км², що може бути результатом вирубки лісів. Перетворення водних об'єктів у відкриті ґрунти і рослинність є наслідком підриву Каховської ГЕС.

На 45,71 км² зменшилися площі зайняті водою, вони перетворилися на пісок – 25,04 км² та відкриті ґрунти – 20,67 км². Це досить великий показник, який підтверджує масштаби витоку води після руйнації дамби всередині.



Рис. 3.30. Гістограма площ виявлених змін землекористувань за 20.05-01.08.2023 рр. (частина 1)

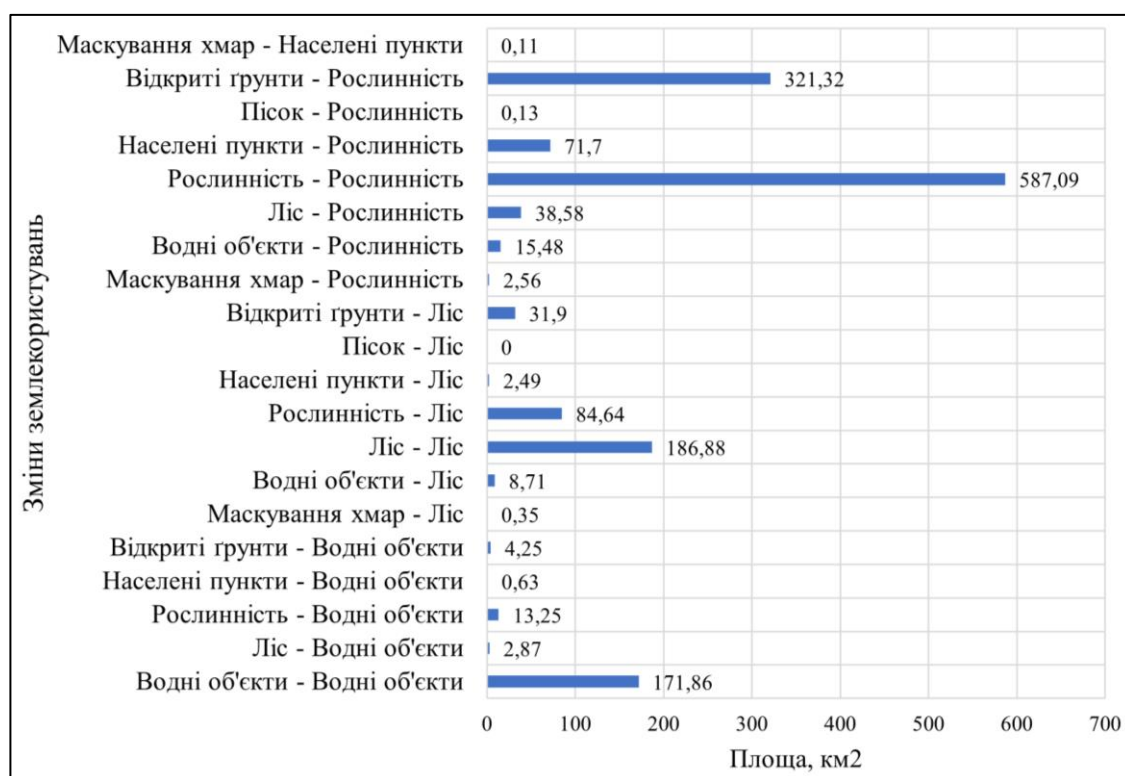


Рис. 3.31. Гістограма площ виявлених змін землекористувань за 20.05-01.08.2023 рр. (частина 2)

Для кращого сприйняття дані було візуалізовано (рис. 3.32). На рисунку спеціально виділені 2 зони: Херсон та Олешки, місце частини колишнього Каховського водосховища. Адже ці місця зазнали значних змін після катастрофи, перші було затоплено водою, а друге майже повернулось до вигляду до 1956 року.

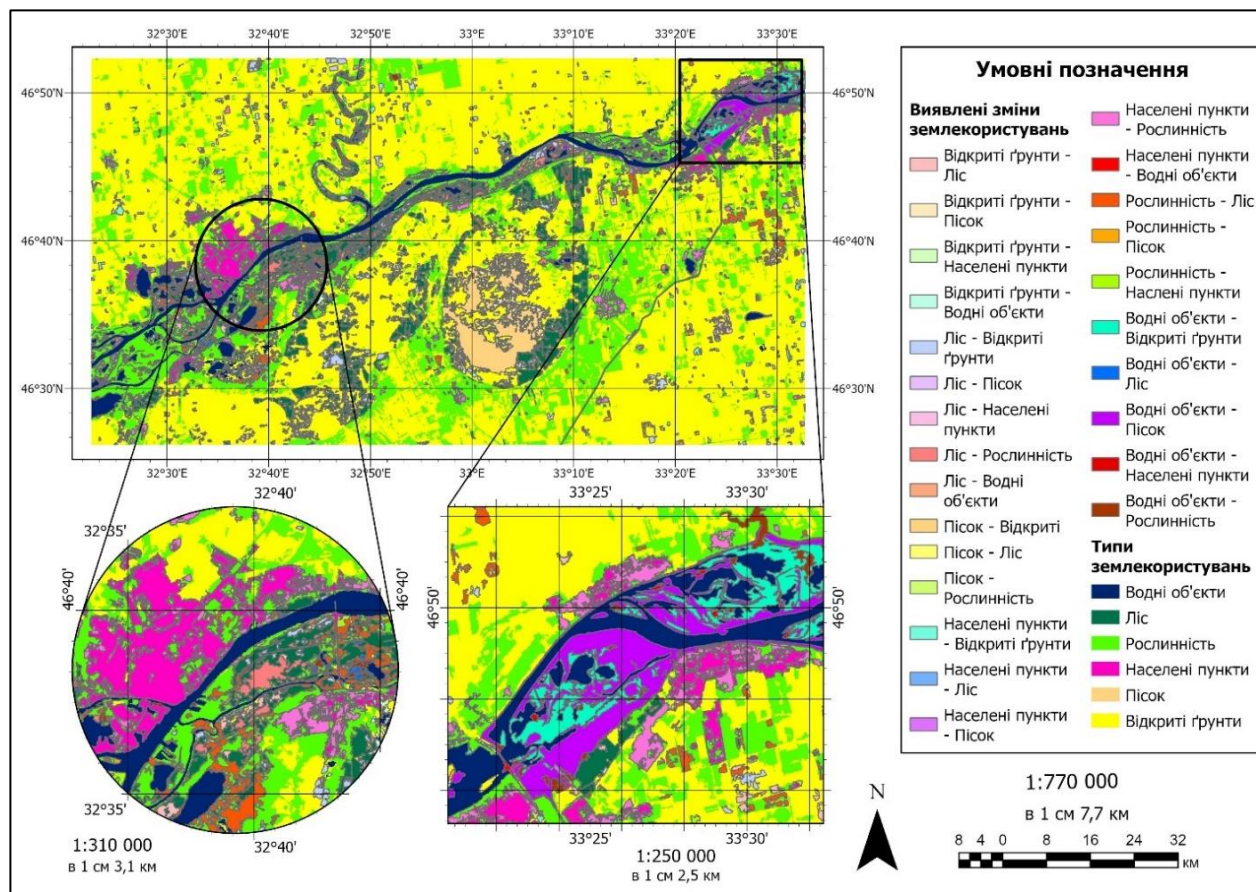


Рис. 3.32. Зміни типів ландшафтів (20.05-01.08.2023 рр.)

За таким же алгоритмом пораховані площі і для комбінацій 2 і 3 (рис. 3.33, 3.34). Найбільше збереглися відкриті ґрунти (1164,86 км²) та рослинність (738,47 км²). Найбільших змін зазнали ті самі відкриті ґрунти та рослинність на 1322,42 км² та 242,05 км² відповідно. Причини були описані вище.

Спостерігається швидка динаміка зменшення лісових насаджень, 126, 89 км² лісу перетворилось на рослинність, що менше за 111,49 км², які залишились без змін. Однак на 16,3 км² (раніше було рослинністю) за цей період додалося хвойних дерев. Можна сказати, що половина залишилась, а половина зникла.

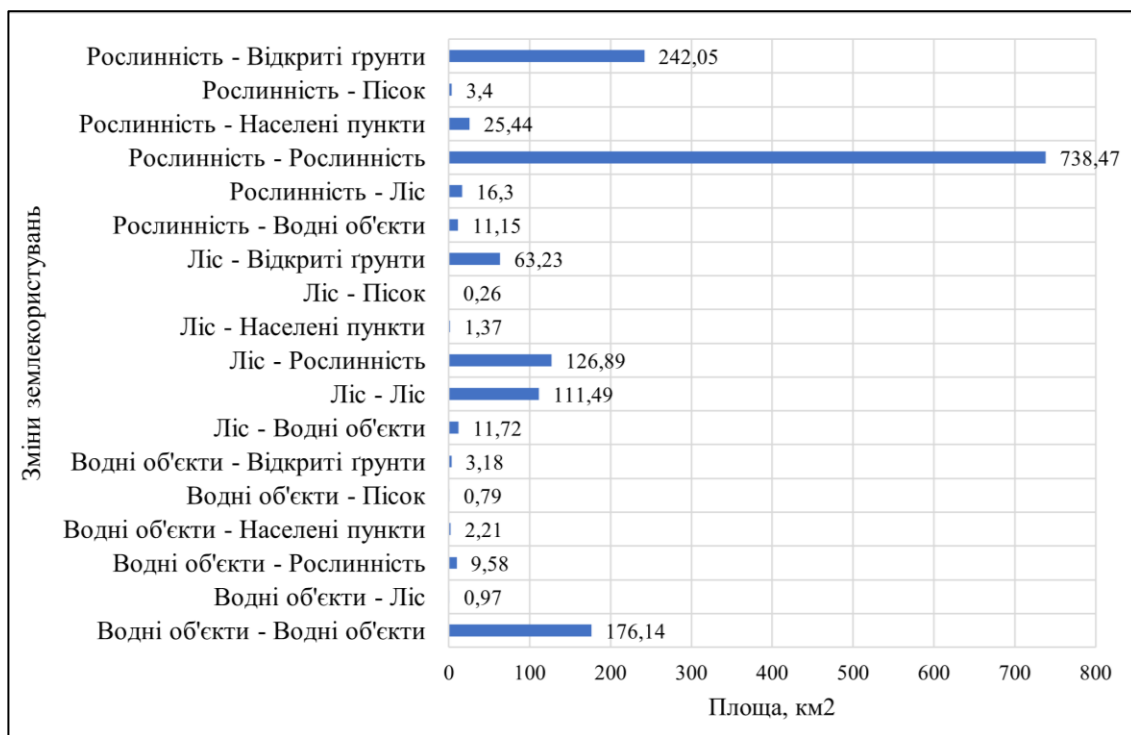


Рис. 3.33. Гістограма площ виявлених змін землекористувань за 01.08.2023-25.05.2024 рр. (частина 1)



Рис. 3.34. Гістограма площ виявлених змін землекористувань за 01.08.2023-25.05.2024 рр. (частина 2)

Візуалізацію отриманих результатів можна побачити на рис. 3.35.

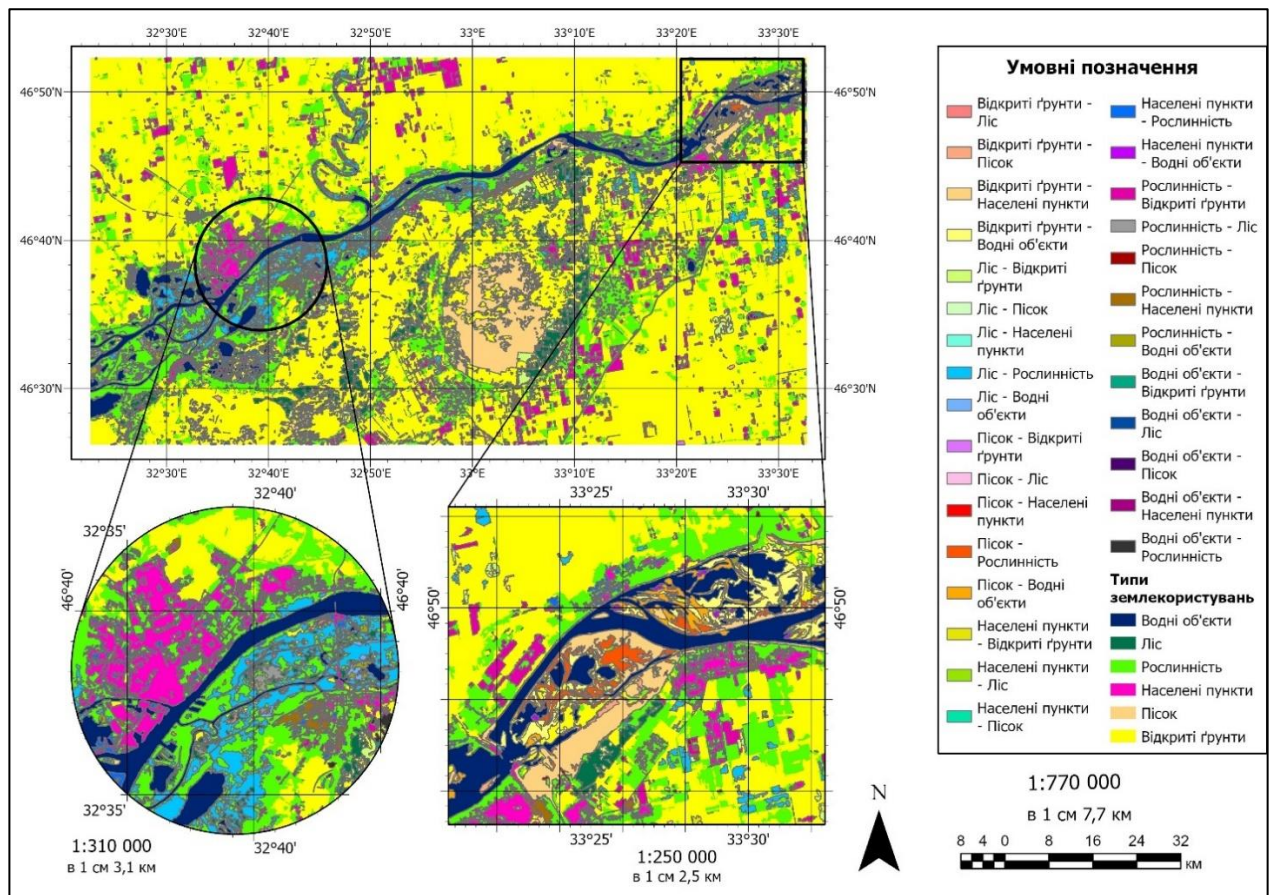


Рис. 3.35. Зміни типів ландшафтів (01.08.2023-25.05.2024 рр.)

Найбільше перетворень у обох випадках видно на тій території, де довгий час була вода, яка більш всього призвела до розмивання ґрунтів на тих ділянках, що може бути причиною такого різкого зменшення лісового покриву. І, звичайно, на місці Каховської ГЕС та водосховища.

На рис. 3.36 наведено порівняння змін однотипних землекористувань, наприклад, перетворення водних об'єктів на водні об'єкти. А на рис. 3.37, 3.38 порівняння різних типів ландшафтів, наприклад, перетворення лісу на пісок.

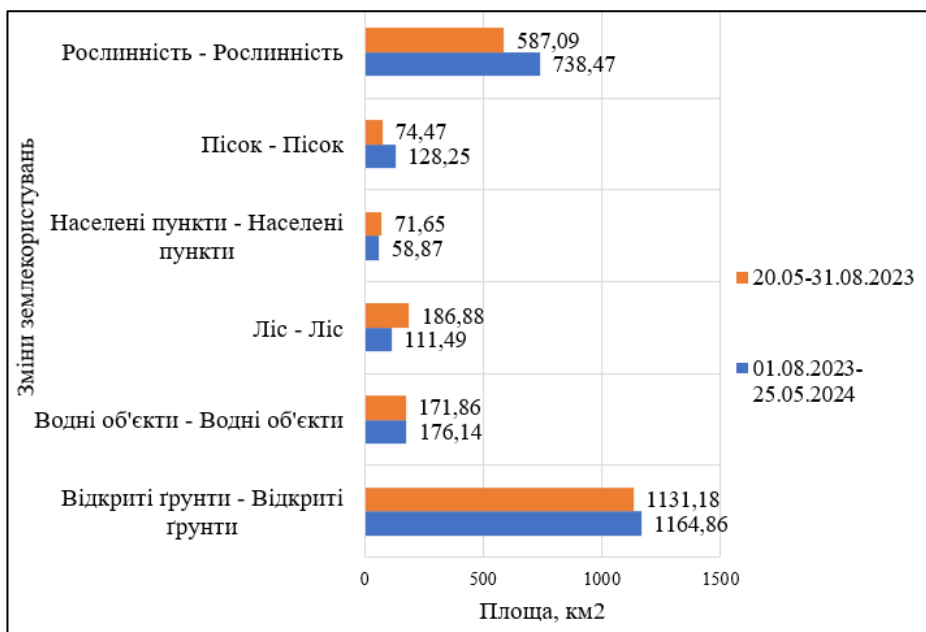


Рис. 3.36. Гістограма площ виявлених змін однотипних землекористувань за 20.05.2023-25.05.2024 рр.

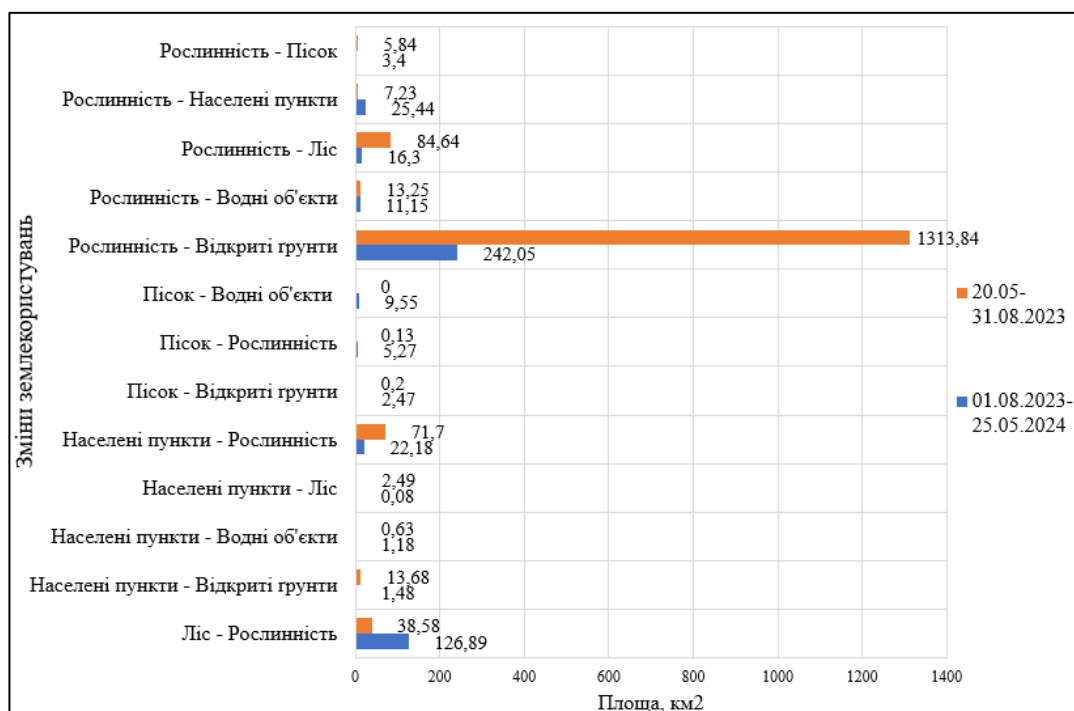


Рис. 3.37. Гістограма площ виявлених змін землекористувань за 20.05.2023-25.05.2024 рр. (частина 1)

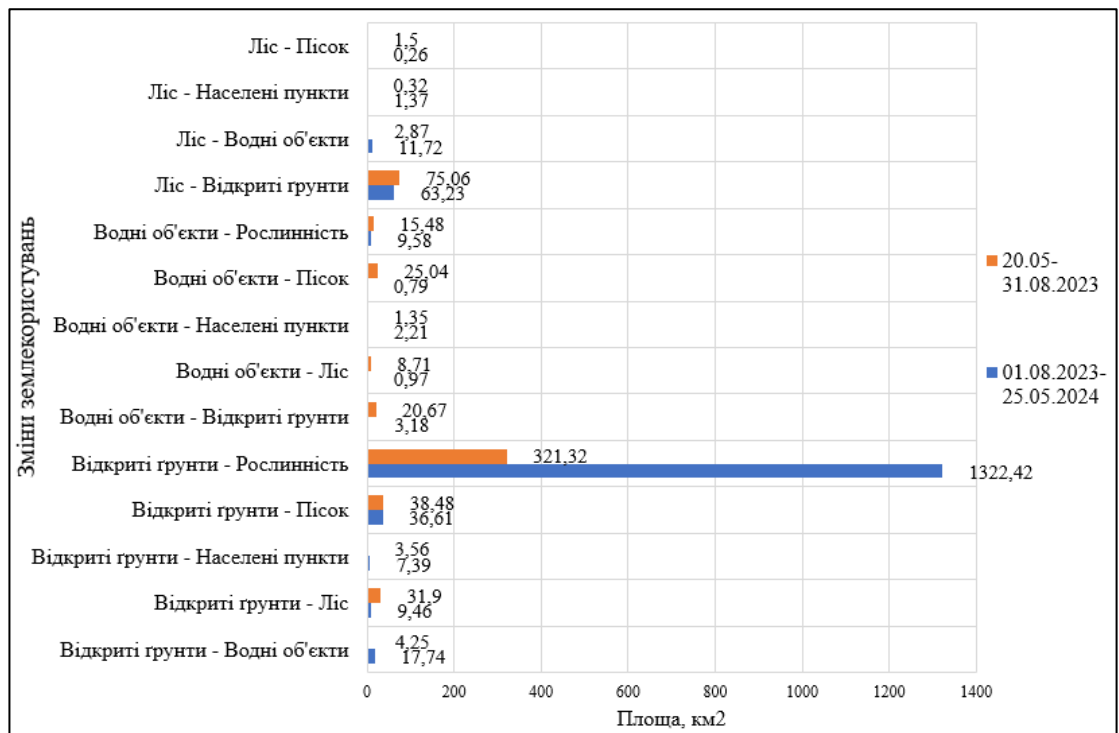


Рис. 3.38. Гістограма площ виявлених змін землекористувань за 20.05.2023-25.05.2024 рр. (частина 2)

ВИСНОВКИ

З використанням платформи хмарних обчислень GEE виконано обробку знімків Sentinel-1 і Sentinel-2 на територію Каховської ГЕС та прибережних населених пунктів Дніпра. Виконано класифікацію знімків методом Random Forest, що дозволило отримати наступні результати щодо змін площ типів ґрунтового покриву.

Площі зайняті рослинністю, відкритими ґрунтами та піском значно збільшили свої площі на 151,38 км², 33,68 км², 53,78 км² відповідно. Порівняно із 2023 роком збільшились на 4,28 км² площі водних об'єктів, що може говорити про збільшення рівня вод Дніпра в даному регіоні.

Значно зменшились площі лісових насаджень на 75,39 км², що може бути результатом великої вирубки лісів. Здебільшого зменшення лісу і водночас збільшення піску спостерігається в одному регіоні – Національному природному парку «Олешківські піски». Також велика частка хвойних дерев зникла саме там, де довгий час земля була покрита водою.

Найбільше змінилися відкриті ґрунти на рослинність (стало більше на 1001,1 км²) та рослинності на відкриті ґрунти (стало менше на 1 071,79 км²).

Значна частина лісу перетворилася у звичайну рослинність (стало більше на 88,31 км²), трохи зменшились перетворення у відкриті ґрунти на 11, 83 км² (75,06 за 2023 р. і 63,23 за 2024 р.). Зменшились перетворення відкритих ґрунтів на пісок. Інші перетворення між типами землекористувань незначні.

Загальна точність класифікації для трьох комбінацій зображень склала 99,41 %, 99,64 % та 99,63 % відповідно. Оцінка Каппа для трьох комбінацій дорівнює 0,99, 0,97 та 0,99, що свідчить про високу узгодженість моделі. Найбільше помилок в усіх випадках виникало при ідентифікації населених пунктів, іноді рослинність і пісок відображалися як міста/села. В усіх випадках правильно відображалися водні об'єкти та рослинність, іноді траплялися похибки із лісом та піском.

Таким чином, отримані результати можна вважати точними та використовувати для подальшого моніторингу змін ландшафтів, планування та управління територіями, зокрема для відновлення постраждалих екосистем та запобігання подібним катастрофам у майбутньому.

СПИСОК ВИКОРИСТАНИХ ДЖЕРЕЛ

1. Gleick P., Vyshnevskiy V., Shevchuk S. Rivers and water systems as weapons and casualties of the russia-ukraine war. *Earth's future*. 2023. Vol. 10, № 11. 13 p. URL: <https://doi.org/10.1029/2023ef003910> (дата звернення: 20.05.2024).
2. The destruction of the Kakhovka dam and its consequences / V. Vyshnevskiy та ін. *Water international*. 2023. Vol. 48, № 5. P. 631-647. URL: <https://doi.org/10.1080/02508060.2023.2247679> (дата звернення: 20.05.2024).
3. Каховська ГЕС: історія та трагедія споруди. *Центр політичного консалтингу* : веб-сайт. URL: <https://cpc.com.ua/articles/kakhovska-ges-istoriya-ta-tragediya-sporudi> (дата звернення: 20.05.2024)
4. Вишневецький В.І. Ріка Дніпро : наукове видання. Київ : Інтерпрес ЛТД, 2011. 384 с.
5. В.І. Вишневецький , В.А. Сташук , А.М. Сакевич. Водогосподарський комплекс у басейні Дніпра : наукове видання. Київ : Інтерпрес ЛТД, 2011. 188 с.
6. Destruction of the kakhovka dam. *Wikipedia, the free encyclopedia* : веб-сайт. URL: https://en.wikipedia.org/wiki/Destruction_of_the_Kakhovka_Dam (дата звернення: 20.05.2024).
7. Geomatic monitoring of environmental hazards in technogenic-loaded territories / S. Nesterenko та ін. *Zemleustriy, kadastr i monìtorìng zemel'*. 2023, № 2. 24 p. URL: <https://doi.org/10.31548/zemleustriy2023.02.03> (дата звернення: 20.05.2024).
8. Rashchenko A., Dyptan T., Malaman A. Activation of man-made processes in the event of a sharp change in the level of reservoirs and groundwater due to man-made disasters of hydraulic structures. *Bases and foundations*. 2023. № 46. P. 123–132. URL: <https://doi.org/10.32347/0475-1132.46.2023.123-132> (дата звернення: 20.05.2024).

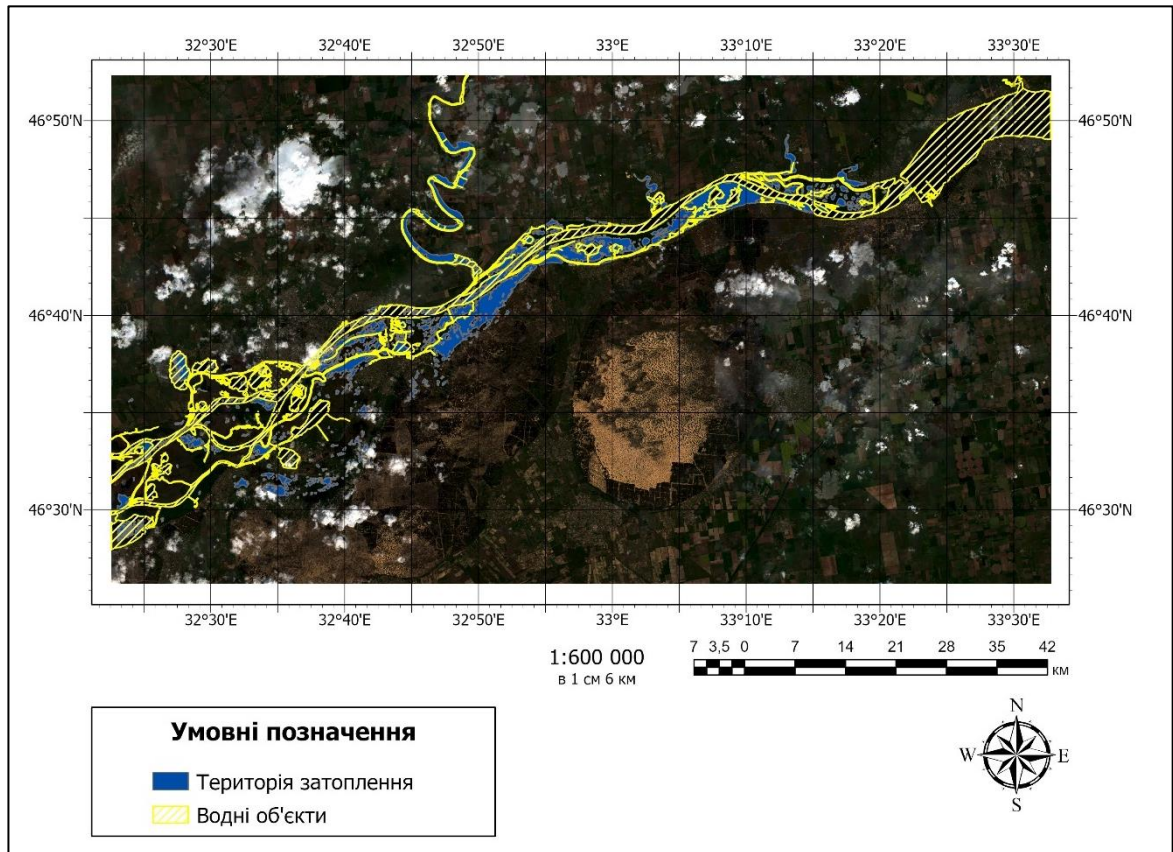
9. Why the evidence suggests russia blew up the kakhovka dam. *The New York Times* : веб-сайт. URL: <https://www.nytimes.com/interactive/2023/06/16/world/europe/ukraine-kakhovka-dam-collapse.html> (дата звернення: 20.05.2024).
10. Before and after satellite images show profound toll of Ukraine dam collapse. *AP News* : веб-сайт. URL: <https://apnews.com/article/ukraine-russia-dam-collapse-before-after-photos-b504eb2ce21e2c30cbcf902fbd718b71> (дата звернення: 20.05.2024)
11. Environmental consequences of the explosion of the kahovsky hydroelectric plant on biodiversity / О. Nepsha та ін. *Revista de la universidad del zulia*. 2023. Vol. 42, № 15. P. 163–178. URL: <https://doi.org/10.46925//rdluz.42.09> (дата звернення: 20.05.2024).
12. The consequences of the russian terrorist attack on the kakhovka hydroelectric power plant (HPP) for wildlife. *Ukrainian Nature Conservation Group* : веб-сайт. URL: <https://uncg.org.ua/en/the-consequences-of-the-russian-terrorist-attack-on-the-kakhovka-hydroelectric-power-station-hps-forwildlife/> (дата звернення: 20.05.2024).
13. Kutz J. N., Brunton S. L. Data-Driven science and engineering: machine learning, dynamical systems, and control. Cambridge : University Press, 2019. 472 p.
14. Segal M. R. Machine learning benchmarks and random forest regression. UC San Francisco : Center for Bioinformatics and Molecular Biostatistics, 2004. 16 p. URL: <https://escholarship.org/uc/item/35x3v9t4> (дата звернення: 21.05.2024).
15. Breiman, L. Random Forests. *Machine Learning*. 2001. P. 5–32. URL: <https://doi.org/10.1023/A:1010933404324> (дата звернення: 21.05.2024).
16. Робастне керування. *Вікіпедія*: веб-сайт. URL: https://uk.wikipedia.org/wiki/Робастне_керування (дата звернення: 21.05.2024).
17. James G., Hastie T., Witten D. An introduction to statistical learning: with applications in R. Springer, 2013. 442 p.

18. Статистичний бутстреп. *Вікіпедія* : веб-сайт. URL: https://uk.wikipedia.org/wiki/Статистичний_бутстреп (дата звернення: 21.05.2024).
19. ВЧ-розмірність . *Вікіпедія* : веб-сайт. URL: <https://uk.wikipedia.org/wiki/ВЧ-розмірність> (дата звернення: 21.05.2024).
20. Кохан С. С., Востоков А. Б. Дистанційне зондування Землі: теоретичні основи : підручник. Київ : Вища школа, 2009. 526 с.
21. Machine learning based combinatorial analysis for land use and land cover assessment in Kyiv City (Ukraine) / V. Belenok та ін. *Journal of applied remote sensing*. 2023. Vol. 17, № 1. 24 p. URL: <https://doi.org/10.1117/1.jrs.17.014506> (дата звернення: 22.05.2024).
22. Code Earth Engine. *Code Earth Engine* : веб-сайт. URL: <https://code.earthengine.google.com/> (дата звернення: 22.05.2024).
23. ArcGIS Pro. *GIS Mapping Software, Location Intelligence & Spatial Analytics* : веб-сайт. URL: <https://www.esri.com/enus/arcgis/products/arcgis-pro/overview> (дата звернення: 23.05.2024).
24. Region Group (Spatial Analyst)–ArcMap. *Documentation* : веб-сайт. URL: <https://desktop.arcgis.com/en/arcmap/latest/tools/spatial-analyst-toolbox/region-group.htm> (дата звернення: 25.05.2024).
25. Set Null (Spatial Analyst)–ArcMap. *Documentation* : веб-сайт. URL: <https://desktop.arcgis.com/en/arcmap/latest/tools/spatial-analyst-toolbox/set-null.htm> (дата звернення: 29.05.2024).
26. Nibble (Spatial Analyst)–ArcMap. *Documentation* : веб-сайт. URL: <https://desktop.arcgis.com/en/arcmap/latest/tools/spatial-analyst-toolbox/nibble.htm> (дата звернення: 29.05.2024).
27. Assessing positional and thematic accuracies of maps generated from remotely sensed data. *Remote sensing handbook - three volume set*. 2018. P. 583–601. URL: <https://doi.org/10.1201/b19355-47> (дата звернення: 23.05.2024).

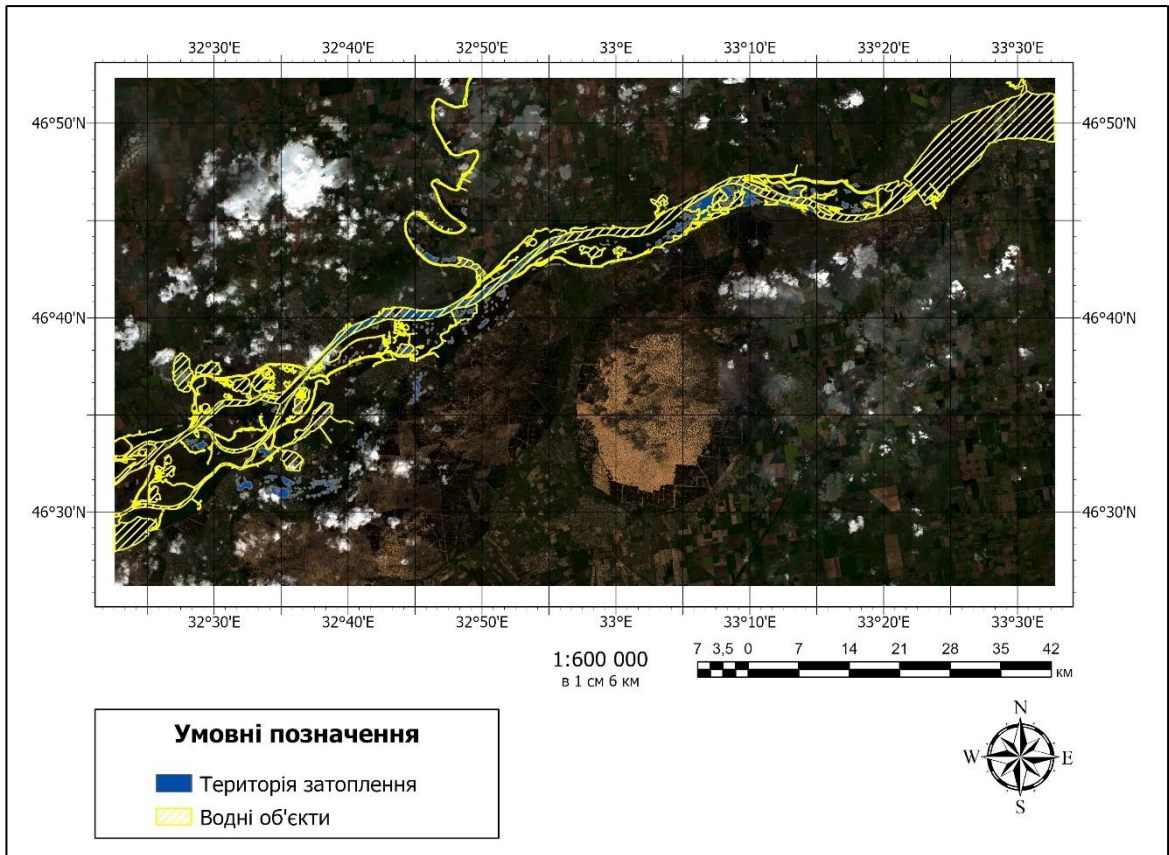
28. S1 mission. *SentiWiki Home* : веб-сайт. URL: <https://sentiwiki.copernicus.eu/web/s1-mission#Sentinel-1-Mission> (дата звернення: 22.05.2024).

29. S2 mission. *SentiWiki Home* : веб-сайт. URL: <https://sentiwiki.copernicus.eu/web/s2-mission> (дата звернення: 22.05.2024).

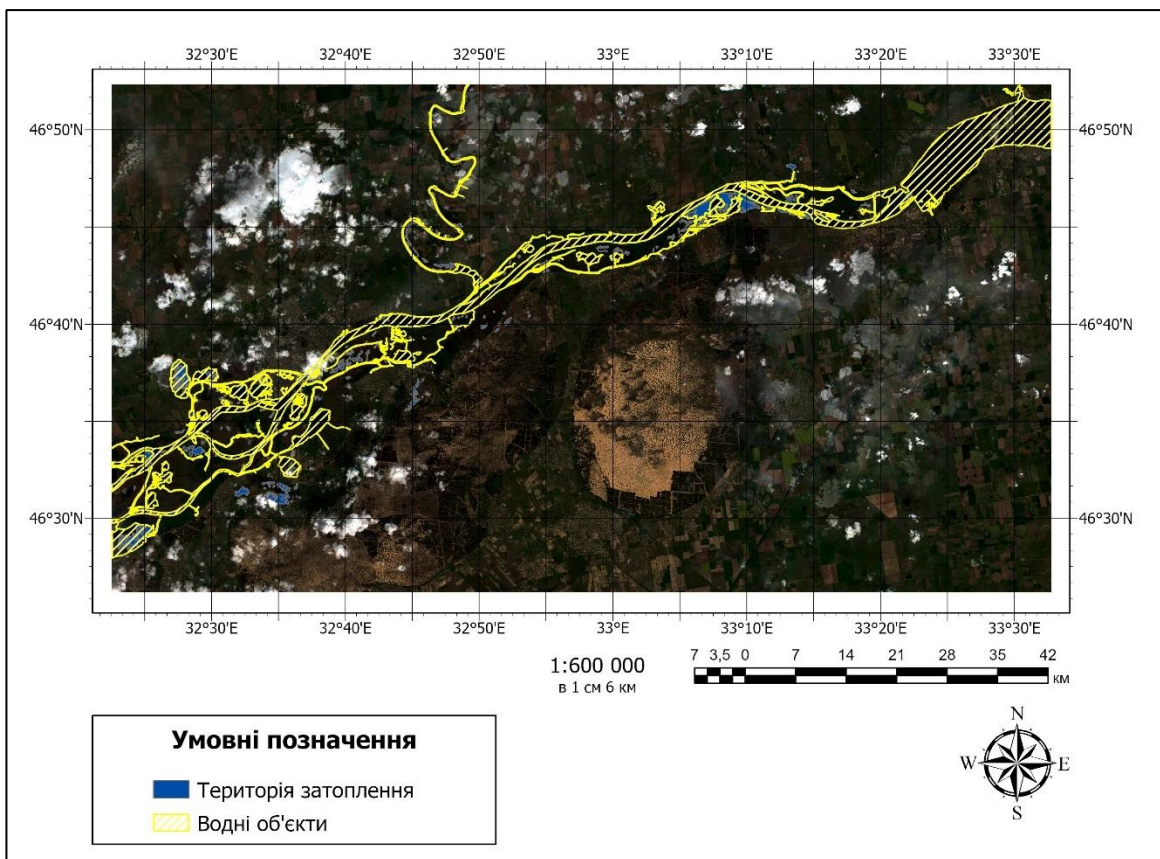
Затоплені території після підриву за 14.06.2023



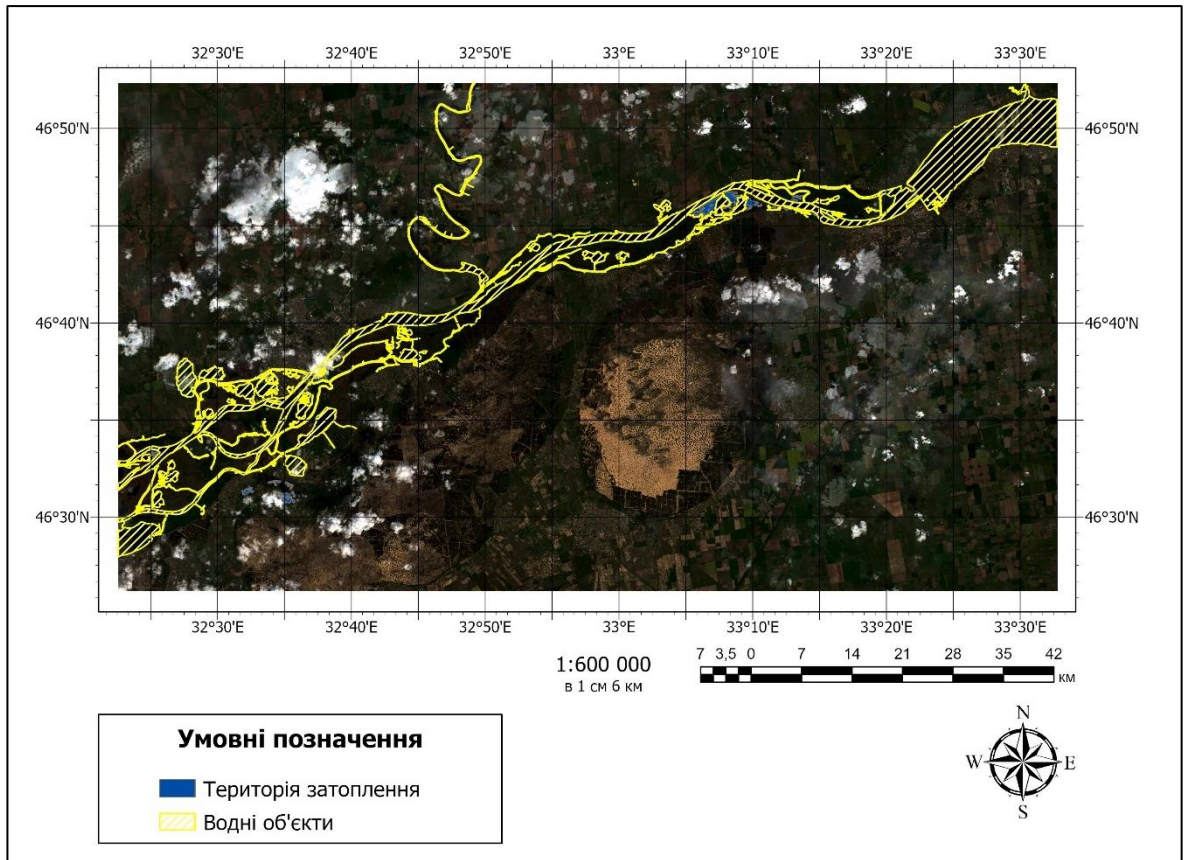
Затоплені території після підриву за 21.06.2023



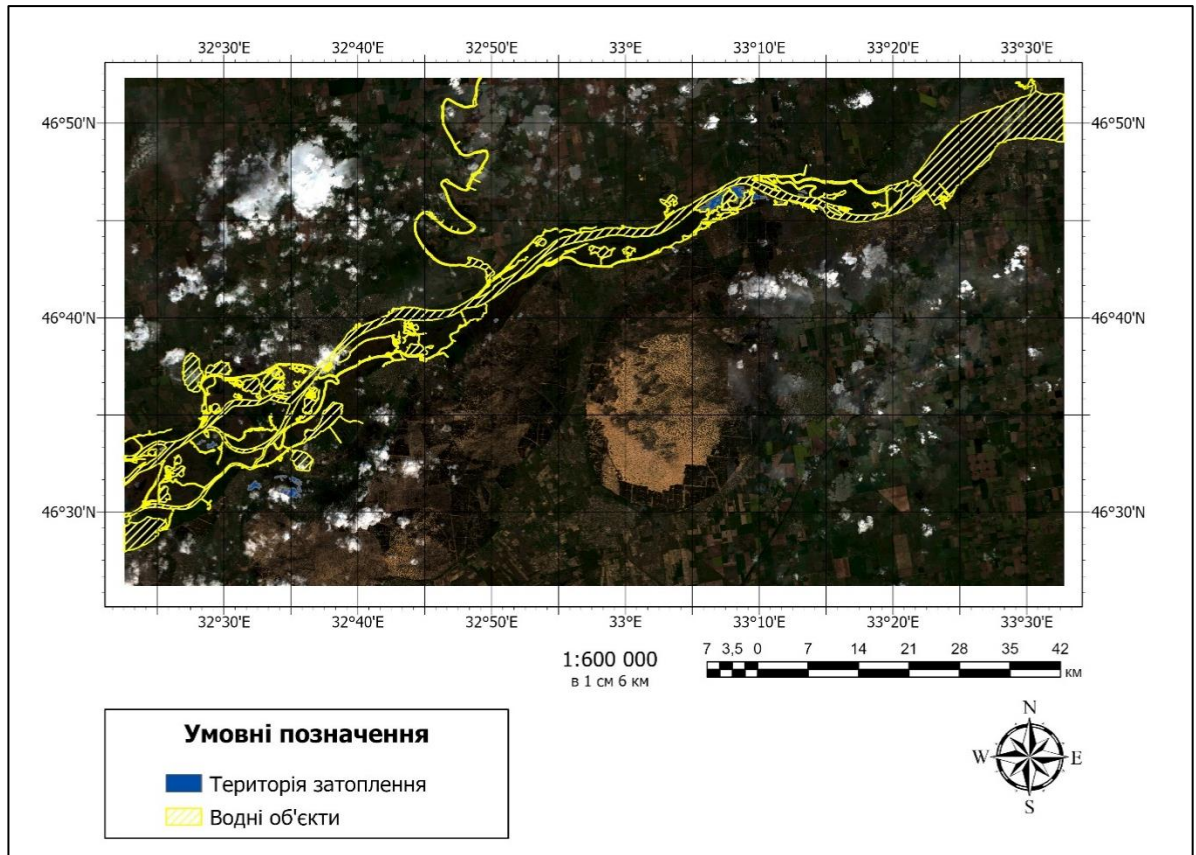
Затоплені території після підриву за 26.06.2023



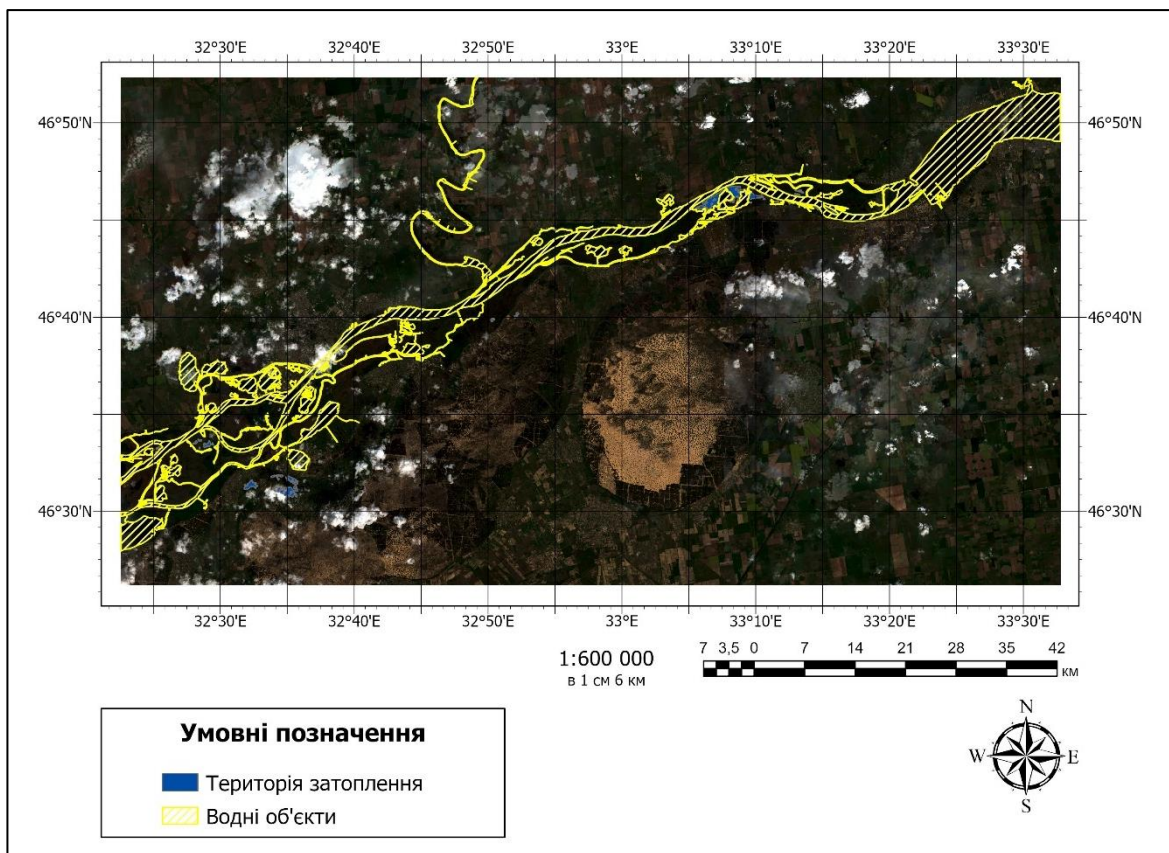
Затоплені території після підриву за 03.07.2023



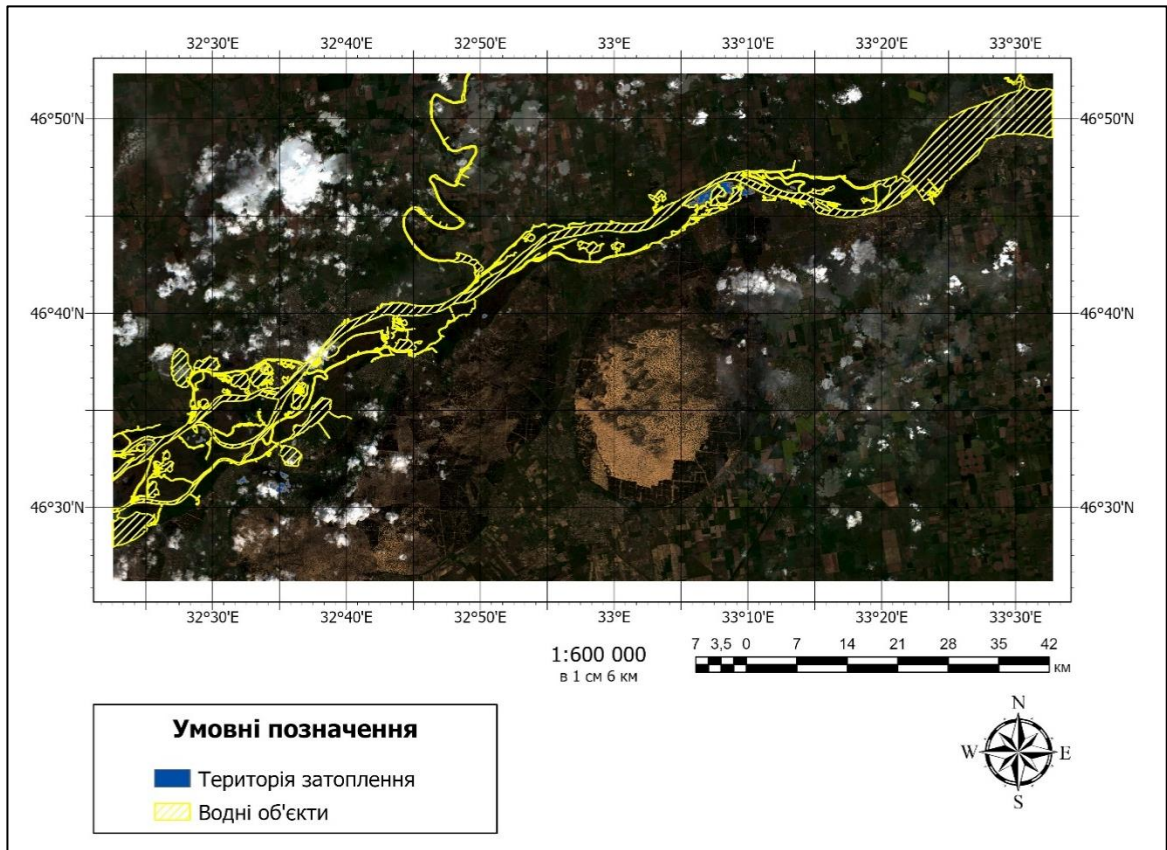
Затоплені території після підриву за 08.07.2023



Затоплені території після підриву за 15.07.2023



Затоплені території після підриву за 20.07.2023



Затоплені території після підриву за 27.07.2023

