

ДОВГОСТРОКОВИЙ МОНІТОРИНГ ЯКОСТІ ПОВЕРХНЕВИХ ВОД ТА ПОТЕНЦІАЛУ ПІДЗЕМНИХ ЗА ДОПОМОГОЮ МЕТОДІВ ОБЧИСЛЮВАЛЬНОГО ІНТЕЛЕКТУ, ГЕОІНФОРМАЦІЙНИХ ТЕХНОЛОГІЙ ТА ДИСТАНЦІЙНОГО ЗОНДУВАННЯ ЗЕМЛІ

Клімов С. В., Старовойт Т. В.

Анотація. Дефіцит води та погіршення її якості, спричинені швидким зростанням населення, урбанізацією та інтенсивною промисловістю підкреслюють важливість ефективного управління водними ресурсами. Враховуючи останні досягнення в дистанційному зондуванні землі, хмарних обчисленнях і обчислювальному інтелекті, вкрай важливо використовувати сучасні передові технології для моніторингу та оцінки якості поверхневих вод в контексті управління водними ресурсами. Метою даного дослідження є розробка гібридних інтелектуальних моделей для оцінки якості поверхневих вод та потенціалу підземних. Гібридні моделі були розроблені з використанням наборів даних зображень Landsat та даних з Водної інформаційної системи Європи (WISE) з застосуванням гібридних нейронних мереж глибокого навчання. Кореляційний аналіз виявив сильний зв'язок між даними дистанційного зондування Землі та різними параметрами, такими як: хлорофіл-а, розчинений кисень, загальний азот та загальний фосфор. Було визначено оптимальні комбінації діапазонів для кожного параметра, що значно підвищило точність оцінки якості води. Моделі на основі гібридних нейронних мереж глибокого навчання продемонстрували вищу точність в порівнянні з моделями, які базувались на класичних методах машинного навчання. Практичність розроблених моделей була успішно перевірена на реальних сценаріях, підтвердивши її як оперативний інструмент управління водними ресурсами.

Ключові слова: обчислювальний інтелект, нечітка логіка, дистанційне зондування землі, супутникові знімки, моніторинг якості поверхневих вод, оцінка потенціалу підземних вод, гібридні нейронні мережі, NEFCLASS-EM, TS-FNN, Fuzzy C-Means, K-Means.

1. ВСТУП

Здоров'я людини, продовольча безпека, економічне зростання, виробництво енергії та екосистеми залежать від води. Зростання населення та урбанізація, інтенсивний розвиток промисловості, збільшення попиту та неправильне використання води посилили водну проблему, зробивши воду дефіцитним і дорогим ресурсом, особливо в країнах що розвиваються. Дана проблема була визнана та було розроблено декілька стратегій до 2030 року на поліпшення ситуації з запобіганням подальшого зниження якості та кількості води [4]. У Європі Водна рамкова директива (ВРД) [5] прямує до досягнення принаймні доброго стану водних ресурсів. Для оцінки стану необхідно проводити моніторинг біологічних, гідроморфологічних та фізико-хімічних показників якості води. Згідно даної директиви річки з площею водозбірного басейну понад 10 км² та озера з площею понад 0,5 км² повинні бути включені до оцінки та моніторингу стану води [5].

Параметри якості води традиційно визначаються шляхом збору зразків на місці та подальшому їх аналізі в лабораторії [4]. Даний метод забезпечує високу точність, але й є трудомістким та затратним за часом та фінансуванням, отже вимагає значних фінансових інвестицій. Крім того, традиційний метод визначає концентрацію необхідних показників лише у точці відбору проб. Тоді як якість води у водоймах рідко є постійною через непередбачуваність подій, такі як випадкові або навмисні витoki з промислових об'єктів і інші чинники. Через це точний моніторинг якості води є складним завданням.

Щоб подолати дані обмеження ми використали технологію дистанційного зондування Землі (ДЗЗ), яка має перевагу великого просторового охоплення та високої часової та роздільної здатності, яка була використана для ідентифікації та моніторингу водних ресурсів більш ефективно та результативно [6, 7, 8]. Дистанційний моніторинг показників якості води базується на встановленні кореляції між даними моніторингу та відповідним відображенням поверхні. Спектральні характеристики води є функціями гідрологічних, біологічних та хімічних характеристик води [9]. І саме, кількість випромінювання хвиль на різних довжинах, відбитих від поверхні води, може прямо чи опосередковано використовуватись для виявлення показників якості води [10].

Чиста вода має здатність відбивати світло з довжиною хвилі більше за 600 нм, що забезпечує високий коефіцієнт відбиття синьо-зеленого кольору, одночасно поглинаючи випромінювання в ближній інфрачервоній (NIR) спектра та за її межами. Збільшення концентрації хлорофілу збільшує поглинання червоного (R) та сильно поглинає синє (B) світло, тоді як пік відбиття розташований у зеленій (G) частині спектра [11]. Прозорість води залежить від загальної концентрації завислих речовин. Дана концентрація є мірою ваги неорганічних частинок, зважених у товщі води та відповідає за більшу частину розсіювання. Впливаючи на розсіювання світла, концентрація завислих речовин (КЗР) у воді безпосередньо контролює прозорість і вміст кисню у водоймі [12]. Підвищена концентрація КЗР спричиняє зміщення піку від G до R-області та збільшує відбиття води в NIR-області.

Таким чином, значна кількість досліджень використовували комбінації смуг та спектральні індекси для розробки емпіричних алгоритмів для оцінки оптично-активних показників якості води та досягли хороших результатів [13 - 14].

Оскільки, зв'язок між поверхневим відбиттям та концентрацією параметрів оцінки якості води є непрямим та нелінійним, їх оцінка є проблематичною, особливо якщо дана оцінка базується на традиційних емпіричних алгоритмах. Протягом останнього десятиліття, з величезним зростанням обчислювальної потужності та розвитком штучного інтелекту і алгоритмів машинного навчання (ML), дані технології все частіше використовуються для вирішення даної проблеми. Найпоширенішими моделями машинного навчання, які використовують в задачах оцінки якості води, є Random Forest (RF), Supported Vector Machine (SVM) і штучна нейронна мережа (ANN).

Гуо та ін. [15] використовували відображення Landsat і MODIS і SVM для моніторингу DO в озері Гурон. Результати демонструють хорошу міцність із середнім $R^2 = 0,91$. Qian та ін. [16] протестували множинну лінійну регресію (MLR), SVM, RF та ANN для моніторингу трьох неоптичних (pH, DO, електропровідність (EC)) та одного оптичного параметра (каламутність) у водосховищі Qingcaosha на основі зображень Sentinel 2. Результати показали, що ANN показав більш надійну продуктивність для всіх WQP (RMSE: 0,33; 0,49; 0,38; 0,26 для pH, DO, EC та Turbidity відповідно) порівняно з традиційними алгоритмами ML.

Гуо та ін. [17] відстежували TP, TN і хімічну потребу в кисні (COD) за допомогою зображень Sentinel 2 і алгоритмів NN, RF і SVM. Їх результати показали, що ML може значно підвищити точність оцінки неоптичних параметрів за допомогою нормалізованої середньоквадратичної помилки (NRMSE) TP: 16,8%; TN: 29,64% і COD 18,75. Аналогічно, Ref. [18] перевірили ефективність MLR, SVM і ANN для моніторингу chl-a, DO, каламутності, синьо-зелених водоростей (BGA) і флуоресцентної розчиненої органічної речовини (fDOM) із зображень Sentinel 2 і Landsat 8. DNN перевершив алгоритми ML, що

призвело до середньоквадратичної помилки (RMSE) 0,86, 7,56, 1,81, 14,50 і 5,19 для BGA, chl-a, DO, fDOM і Turbidity відповідно.

В дослідженнях [18, 19, 20, 21] було показано, що ANN і SVM забезпечили чудову продуктивність у моніторингу як оптично активних, так і неактивних показників якості води. Загалом, штучні нейронні мережі, як метод лінійної апроксимації, є більш гнучким для моніторингу показників якості води. Однак результуюча точність моделей машинного навчання зазвичай є функцією вибраної моделі та якості залежної від розміру навчальних даних. Розробка моделей ШНМ вимагає великих навчальних наборів даних та великого досвіду, щоб побудувати оптимальну архітектуру штучних нейронних мереж. Використання занадто великої кількості шарів може призвести до переобладнання, яке передбачає підгонку шуму в навчальних даних [23]. З іншого боку, мала кількість шарів може призвести до недостатнього підбору, коли модель не може належним чином представити складність даних [23].

Метою даного дослідження є розробка гібридних моделей обчислювального інтелекту, а саме нейро-нечітких нейронних мереж на базі просторових даних з застосуванням алгоритмів дистанційного зондування землі та переваг геоінформаційного моделювання, з порівнянням їх ефективності. В даній статті пропонується поєднання нейро-нечітких нейронних мереж з метаевристичними алгоритмами та алгоритмами ДЗЗ, для оцінки якості води по супутниковим знімкам та оцінюється ефективність отриманих результатів. Саме гібридні нейронні мережі в поєднанні з ДЗЗ та ГІС можуть революціонізувати методи оцінки якості води та потенціалу підземних вод. Ці інструменти можуть допомогти точно визначати фактори, які призводять до погіршення якості води та потенційної появи неочікуваних поверхневих утворень у внутрішніх водоймах. Крім того, дане дослідження має на меті вивчити ступінь впливу цих факторів на екологічний стан водойм з часом.

2. МАТЕРІАЛИ ТА МЕТОДИ

2.1. Область дослідження

Областю дослідження (Рис. 1) ми обрали територію міста Києва та Київської області (Україна). Територія дослідження розташована в південно-східній Європі. Більшість річок належить до басейну Чорного моря. Найбільша річка Дніпро, яку ми й обрали для досліджень. Для навчання штучних нейронних мереж та нейро-нечітких нейронних мереж використали дані з Водної інформаційної системи Європи (WISE).

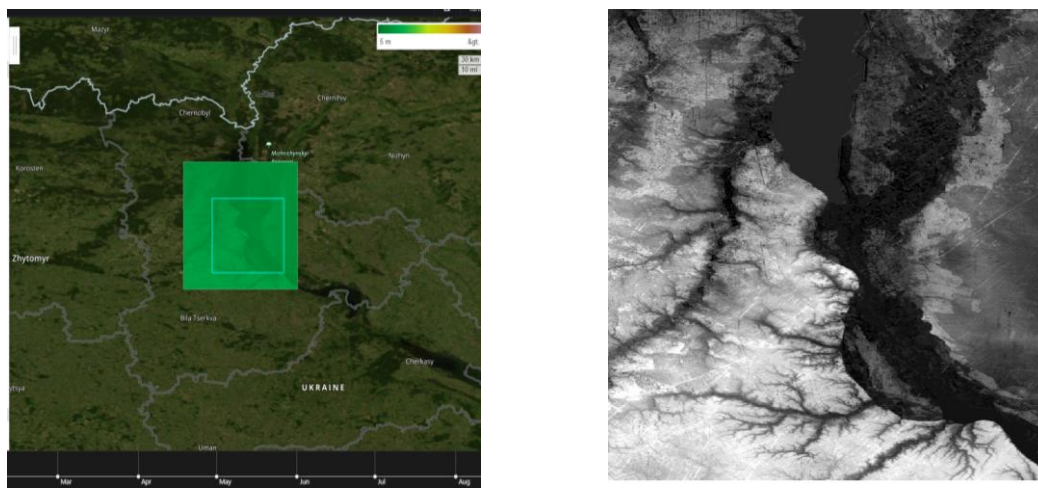


Рисунок 1. – Територія дослідження

2.2. Підготовка даних

Дистанційний моніторинг показників якості води базується на кореляції між вимірюванням на місці та відповідними коефіцієнтами відбиття від поверхні. Для даного дослідження використовувались супутникові знімки поверхні Landsat 5, Landsat 7 і Landsat 8 з 2010 по 2022 рік над Європою. Загалом було проаналізовано більше двох тисяч зображень для створення часових рядів та навчання моделей для моніторингу.

Зображення відбиття від поверхні Landsat скориговано за атмосферою та містить шість (B, G, R, NIR, SWIR, SWIR 2) смуг, оброблених до орторектифікованого відбиття поверхні за допомогою LEDAPS [24]. Супутники Landsat досягають максимального покриття поверхні один раз на 16 днів, з просторовою роздільною здатністю 30 м для мультиспектральних діапазонів. Google Earth Engine API, інтегрований у Google Colab, використовувався як точка доступу до зображень.

Для кожної точки значення коефіцієнта відбиття поверхні були отримані з доступних зображень Landsat Surface Reflectance Level 2A. Маскування хмар і тіней було виконано, щоб забезпечити пікселі чистої води. Отримана таблиця містила ідентифікатор станцій моніторингу, відповідне значення коефіцієнта відбиття поверхні та дату зондування. Коефіцієнт відбиття поверхні був відфільтрований за датою, щоб відповідати даним на місці. Максимальний часовий проміжок між відбором проб на місці та естакадою супутника становив 3 дні.

Кореляційний аналіз Пірсона використовувався для дослідження зв'язку між дистанційним зондуванням і даними на місці з коефіцієнтом кореляції (r). На основі кореляції було визначено набір вхідних даних для кожного показника якості води. Дані були стандартизовані відповідно до нормального розподілу із середнім значенням 0 і стандартним відхиленням 1 і розділені на навчальні та тестові набори (80% і 20%, відповідно).

2.3. Модель нечіткої нейронної мережі

Нечіткі теоретичні системи TS здатні використовувати деякі нечіткі математичні правила для отримання більш складних нелінійних функцій, таким чином забезпечуючи перевагу системи за рахунок зменшення точної кількості нечітких правил при роботі з проблемами, що містять кілька змінних [26]. Нечітка теоретична система TS зазвичай визначається в логічній формі «якщо-тоді», а нечіткий висновок теоретичної системи виражається наступним чином [25]:

$$R^i : \text{If } x_1 \text{ is } A_1^i, x_2 \text{ is } A_2^i, \dots, x_k \text{ is } A_k^i \quad (1)$$

$$\text{Then } y_i = p_0^i + p_1^i x_1 + \dots + p_k^i x_k$$

де A_j^i – нечітка множина нечіткої системи; p_j^i ($j = 1, 2, \dots, k$) – це параметри нечіткої системи; y_i вихідне значення, отримане за нечітким правилом; вхідна частина, тобто *If*, є нечіткою, а вихідна частина, тобто *Then*, є детермінованою частиною. Даний нечіткий висновок вказує на те, що вихід є лінійною комбінацією входів [25].

Припустимо, що для входів $x = [x_1, x_2, \dots, x_k]$ ступінь приналежності кожної вхідної змінної x_j спочатку обчислюється нечітке правило [25]:

$$\mu_{A_j^i} = \exp(-(x_j - c_j^i)^2 / b_j^i) \quad (2)$$

$$(j = 1, 2, \dots, k; i = 1, 2, \dots, n)$$

де c_j^i і b_j^i – центр та ширина функції приналежності, k – вхідний параметр, n – кількість нечітких підмножин [25].

Нечіткі операції виконуються над кожним із зазначених вище ступенів приналежності, а нечіткий оператор використовується як конкатенований мультиплікативний оператор [25]:

$$\omega^i = u_{A_j^1}(x_1) \times u_{A_j^2}(x_2) \times u_{A_j^k}(x_k) \quad (3)$$

$$(i = 1, 2, \dots, n)$$

За результатами нечіткого розрахунку вихідне значення моделі y_i виходить [25]:

$$y = \frac{\sum_{i=1}^n \omega^i (p_0^i + p_1^i x_1 + \dots + p_k^i x_k)}{\sum_{i=1}^n \omega^i} \quad (4)$$

В даному дослідженні нечіткі нейронні мережі були застосовані для прогнозування показників якості води та потенціалу підземних вод. Модель нечіткої нейронної мережі зазвичай має чотири рівні: вхідний рівень, рівень фазифікації, рівень обчислення нечітких правил і вихідний рівень [25]. Вхідні шари моделі мережі пов'язані через вектор x_i , тому кількість вузлів модельної мережі узгоджується з розмірністю вхідного вектора. Рівень фазифікації використовує функцію приналежності рівняння (3) для нечіткості вхідних значень, щоб отримати значення приналежності $u_{A_j^i}$. Значення ω на рівні нечітких обчислень отримується за допомогою рівняння (4) послідовного фазового множення, а потім вихідне значення рівня вихідних даних у цій системі нечіткої моделі отримується за рівнянням (5). Алгоритм навчання нечіткої нейронної мережі виглядає наступним чином [25]:

Крок 1. Похибка розрахунку:
$$e = \frac{1}{2} (y_d - y_c)^2 \quad (5)$$

де очікуваний вихід мережі становить y_c , а похибка між очікуваними результатами та фактичними становить e .

Крок 2. Корекція коефіцієнта:
$$p_j^i(k) = p_j^i(k-1) - \alpha \frac{\partial e}{\partial p_j^i} \quad (6)$$

$$\frac{\partial e}{\partial p_j^i} = (y_d - y_c) \omega^i / \sum_{i=1}^m \omega^i \cdot x_j$$

де p_j^i – коефіцієнт нейронної мережі, α – швидкість навчання мережі, x_j – вхідні параметри мережі, ω^i – неперервний добуток приналежності вхідних параметрів.

Крок 3. Корекція параметрів:
$$c_j^i(k) = c_j^i(k-1) - \beta \frac{\partial e}{\partial c_j^i} \quad (7)$$

$$b_j^i(k) = b_j^i(k-1) - \beta \frac{\partial e}{\partial b_j^i}$$

де b_j^i та c_j^i є шириною та центральним значенням репрезентативної функції приналежності в нечіткому правилі.

2.4. Показники оцінки точності моделей

Для оцінки точності моделей ми використали такі чотири показники: загальна точність (OA), помилка неправильної класифікації (ME), помилка пропуску (OE) та значення ROC-AUC. OA – сума пікселів, правильно класифікованих, як вода, поділений на загальну кількість пікселів води, яка представлена матрицею плутанини. OE – кількість пікселів, які належать воді, але класифікуються як інші типи поверхні, які можна ідентифікувати за стовпцем матриці помилок. Значення ROC-AUC [2] – це площа під кривою відношення чутливості (рівняння 8) до специфічності (рівняння 9). Дане значення коливається в діапазоні від 0,50 до 1. Чим більше значення, тим краща продуктивність. У випадку, коли значення перевищує 0,70 – результат класифікації доступний [1].

$$Sensitivity = \frac{TPR}{TPR + FNR} \quad (8)$$

$$Specificity = \frac{TNR}{TNR + FPR} \quad (9)$$

де TPR (справжній позитивний) і FNR (помилковий негативний) – це пікселі, які правильно і неправильно класифіковані як вода, а TNR (справжній негативний) та FPR (хибний позитивний) – це пікселі, які правильно і неправильно класифіковані як не вода [1].

Для кількісної оцінки якості води можна використовувати різні показники для оцінки ефективності моделі якості води, включаючи коефіцієнт детермінації R^2 (рівняння 10), середню абсолютну похибку (MAE, рівняння 11), корінь середньої квадратичної похибки (RMSE, рівняння 12), середню квадратичну похибку (MSE, рівняння 13), залишкове відхилення передбачення (RPD, рівняння 14) та довірчий інтервал (CI, рівняння 15) [1].

$$R^2 = 1 - \frac{\sum_{i=1}^n (y_i - y'_i)^2}{\sum_{i=1}^n (y'_i - \bar{y})^2} \quad (10)$$

$$MAE = \frac{1}{n} \sum_{i=1}^n |y_i - y'_i| \quad (11)$$

$$RMSE = \sqrt{\frac{1}{n} \sum_{i=1}^n (y_i - y'_i)^2} \quad (12)$$

$$MSE = \frac{1}{n} \sum_{i=1}^n (y_i - y'_i)^2 \quad (13)$$

$$RPD = \sqrt{\frac{\sum_{i=1}^n (y_i - \bar{y})^2}{\sum_{i=1}^n (y_i - y'_i)^2}} \quad (14)$$

$$CI = \left[1 - \frac{\sum_{i=1}^n |y'_i - y_i|}{\sum_{i=1}^n (|y'_i - \bar{y}'| + |y_i - \bar{y}|)} \right] \times \left[1 - \frac{\sum_{i=1}^n (y'_i - y_i)^2}{\sum_{i=1}^n (y'_i - \bar{y}')^2} \right] \quad (15)$$

Слід дотримуватись не надто високих значень R^2 , так як занадто високі значення призводять до надмірної адаптації та недостатньої переносимості моделі, оскільки даний показник чутливий до екстремальних значень [3]. Тому на практиці R^2 часто

використовується в поєднанні з RMSE, RPD, CI та іншими індикаторами, щоб збалансувати точність підгонки та обчислювальну складність [1].

3. РЕЗУЛЬТАТИ МОДЕЛЮВАННЯ

3.1. Аерокосмічне дослідження та геоінформаційне моделювання

Однією з головних переваг супутникових технологій є їх здатність вловлювати світло за межами видимого спектру, який може сприймати людське око. Інфрачервоне світло є одним з типів випромінювання, яке можуть виявляти такі супутники як Sentinel-2. Інфрачервоне випромінювання може отримувати інформацію про температуру поверхні, стан рослинності та навіть атмосферні умови. Наприклад, здорова рослинність відбиває більше інфрачервоного світла ніж нездорова рослинність, або без рослинні поверхні.

Кожне растрове зображення має 1009 пікселів у висоту (приблизно 10 кілометрів), 1014 пікселів у ширину (приблизно 10 кілометрів) та приблизно 4 спектральні смуги: синя, зелена, червона, ближня інфрачервона. Супутник записує інформацію на 11 довжинах хвиль. Ми обрали: синій, зелений, червоний та інфрачервоний спектри.

Близький інфрачервоний діапазон є особливо корисним для аналізу рослинного покриву. В даному поєднанні здорова рослинність виглядає яскраво-червоною, тоді як нерослинні поверхні набувають інших кольорів. Дана техніка підсилює контраст типів ґрунтового покриву, та візуалізує те, що непомітно у видимому світлі. На рисунку 2 вказаний зразок візуалізації де ми чітко розділили воду, рослинність та ґрунт.

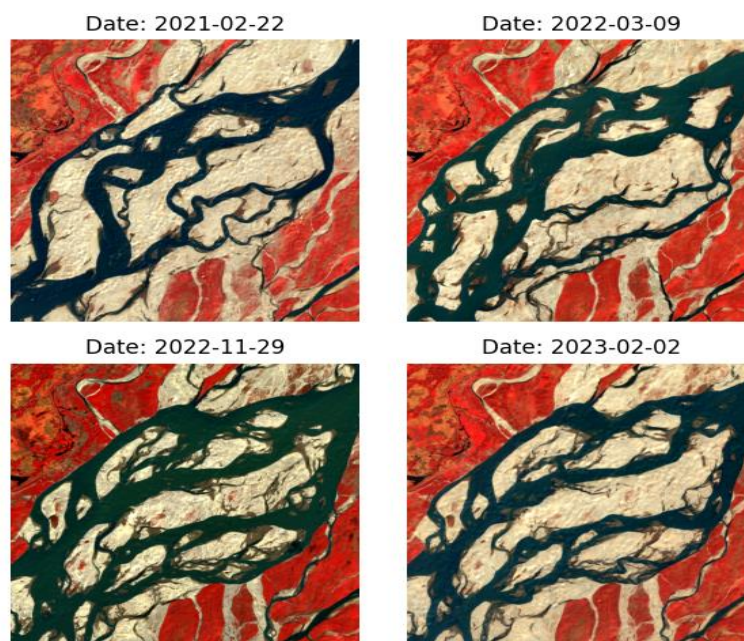


Рисунок 2. – Результат розподілу території на кластери (вода, рослинність та ґрунт)

Після візуалізації зображення в різних колірних просторах, ми дослідили детальніше червоний та інфрачервоний діапазони. Даний крок передбачає створення діаграми розсіювання, де коефіцієнт відбиття червоного пікселя вказаний на осі абсцис, а інфрачервоного – на осі ординат (Рис. 3). Наступним кроком ми перетворили зображення на табличні дані. Кожен рядок отриманої таблиці відповідає одному пікселю для однієї дати. Наприклад, перший рядок – це піксель (1, 1) 22 лютого 2022 року. Кожен стовпець відповідає інтенсивності спектральної смуги (синій, зелений, червоний, інфрачервоний).

Для визначення кластерів, які відповідають водоймам, рослинності та ґрунту ми застосували метод чіткої та нечіткої кластеризації (Рис.3), де один з кластерів відповідає воді. Вода, як правило, має низький коефіцієнт відбиття як в червоному так і в інфрачервоному випромінюванні, тоді як рослинність має високий коефіцієнт відбиття в інфрачервоному випромінюванні, але низький у червоному.

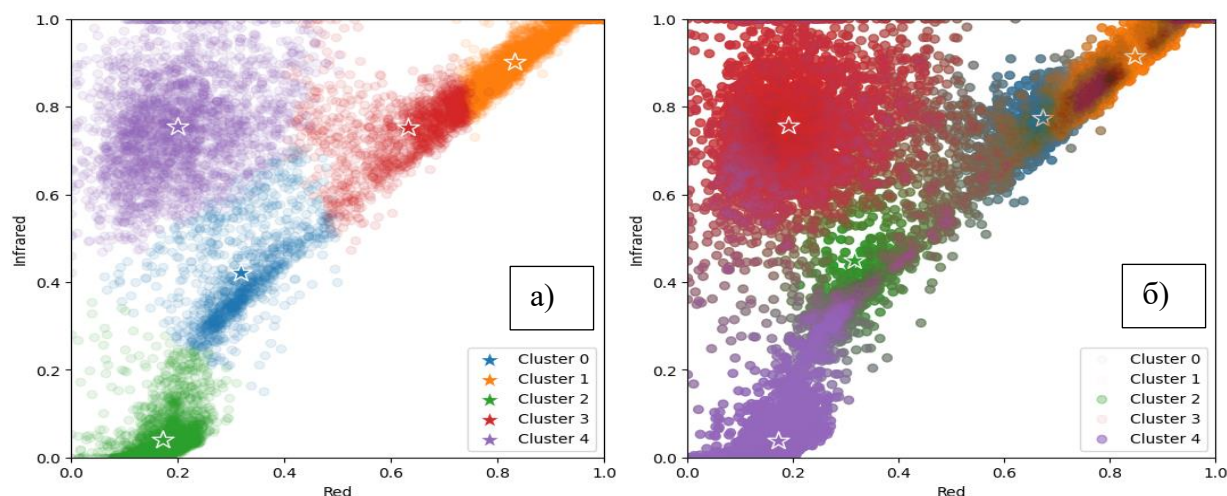


Рисунок 3. – Візуалізація результатів чіткої і нечіткої кластеризації, де а) чітка кластеризація, б) – нечітка кластеризація пікселів.

Для розпізнавання кластерів, які відповідають водоймам, ми застосували такі інтелектуальні алгоритми: K-Means, Fuzzy C-Means (FCM), Нечіткі нейронні мережі (FNN), нечіткі нейронні мережі Такагі-Сугено (TS-FNN), NEFCLASS, NEFCLASS- EM.

3.2. Розробка моделей обчислювального інтелекту

3.2.1. Розробка моделей K-Means та Fuzzy C-Means (FCM)

Перевагою нечіткої кластеризації є те, що кожен піксель отримує оцінку членства для кожного кластера. На рисунку 4 вказаний результат чіткої та нечіткої кластеризації, де вказано ступінь приналежності кожного пікселя до вибраного кластера.

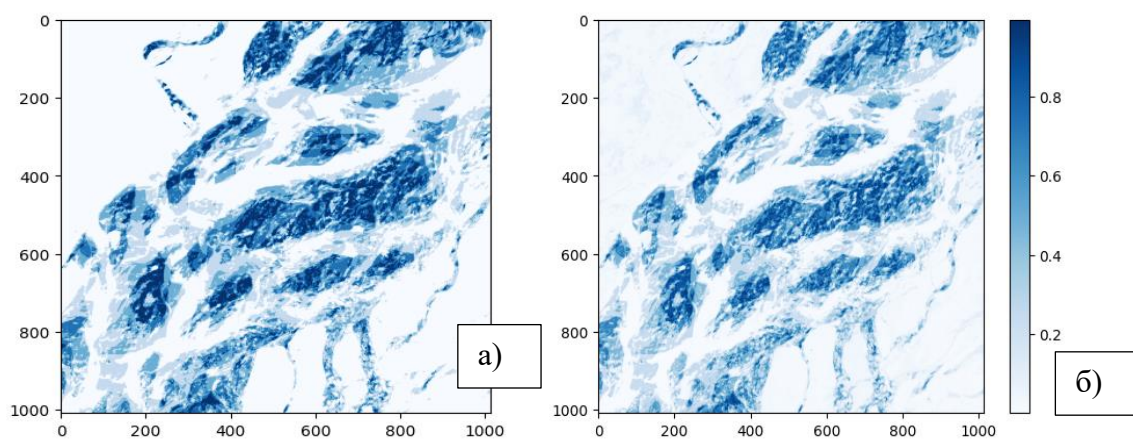


Рисунок 4. – Графічне відображення результатів розпізнавання кластеру «Вода», де а) результат алгоритму K-Means, а б) – це результат алгоритму Fuzzy C-Means

Метод нечіткої кластеризації показує ступінь приналежності кожного пікселя до кожного кластера м'якше в порівнянні з методом жорсткої кластеризації K-Means. Наступним кроком ми дослідили точність нечітких нейронних мереж та гібриду NEFCLASS з алгоритмом електромагнітної евристики.

3.2.2. Модель нечіткої нейронної мережі

Нечіткі нейронні мережі (FNN) – це гібридні моделі, які поєднують нечітку логіку з обчислювальною потужністю нейронних мереж для обробки невизначеностей і неточних даних у таких завданнях, як класифікація, кластеризація та регресія. У контексті супутникових зображень FNN особливо корисні, оскільки супутникові дані часто містять шум, неповну інформацію та властиву нечіткість, особливо коли йдеться про природні явища, такі як хмарність, форми рельєфу, рослинність і водойми.

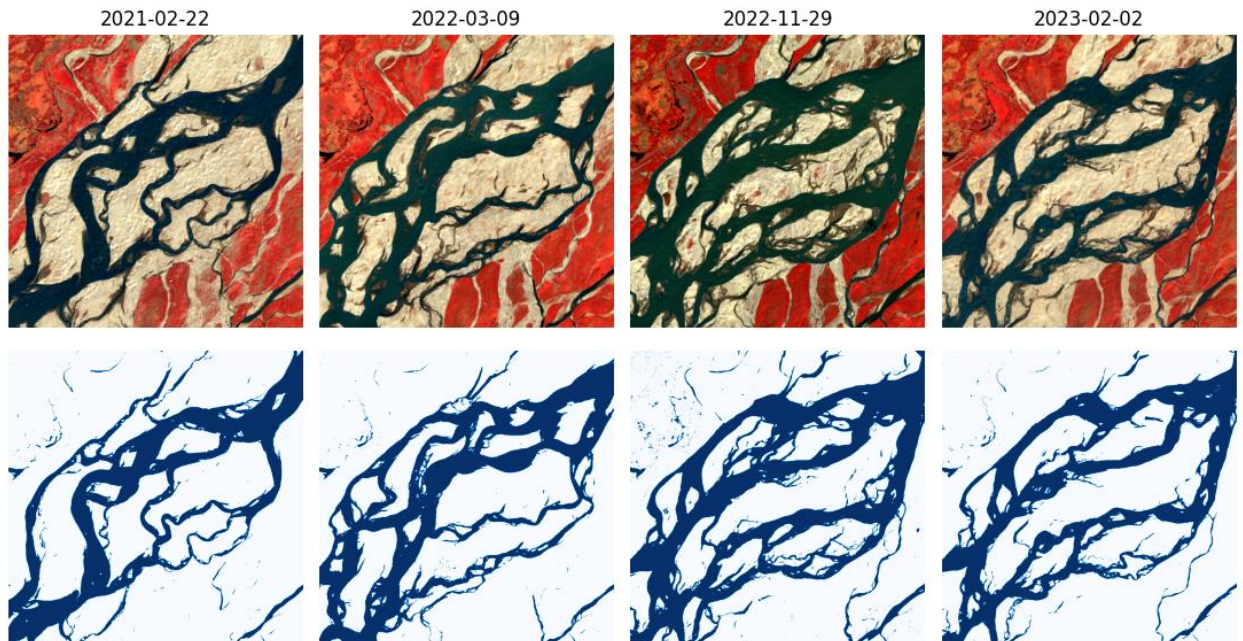


Рисунок 5. – Візуалізація результатів розпізнавання води на супутникових знімках за допомогою нечіткої нейронної мережі (FNN)

На рисунку 5 вказаний результат виокремлення води, де ми бачимо, що нечітка нейронна мережа розпізнала воду на супутникових знімках значно краще, ніж алгоритми K-Means та Fuzzy C-Means. FNN дозволяють використовувати нечіткі значення належності (а не двійкові рішення), для моделювання цих неоднозначних або змішаних пікселів. На відміну від традиційних нейронних мереж, FNN призначають ступені приналежності до різних класів, тому один піксель може належати до кількох категорій з різною ймовірністю або ступенем (наприклад, 70% ліс, 30% вода).

3.2.3. Модель нечіткої нейронної мережі Такагі-Сугено (TS-FNN)

Моделі Такагі-Сугено є типом системи нечіткого логічного висновку, де наслідок (вихід) нечітких правил є або лінійною комбінацією вхідних змінних, або константою. У TS-FNN кожне нечітке правило має відповідну лінійну модель (або константу), яка вивчається, а вихідні дані обчислюються як середньозважене значення вихідних даних правил. Результат реалізації моделі TS-FNN вказаний на рисунку 6. Модель навчена мінімізувати середню квадратичну помилку (MSE) між прогнозованими виходами та базовими мітками істинності. Після навчання мережа прогнозує членство в кластерах для кожного пікселя на супутниковому зображенні, а призначені кластери візуалізуються.

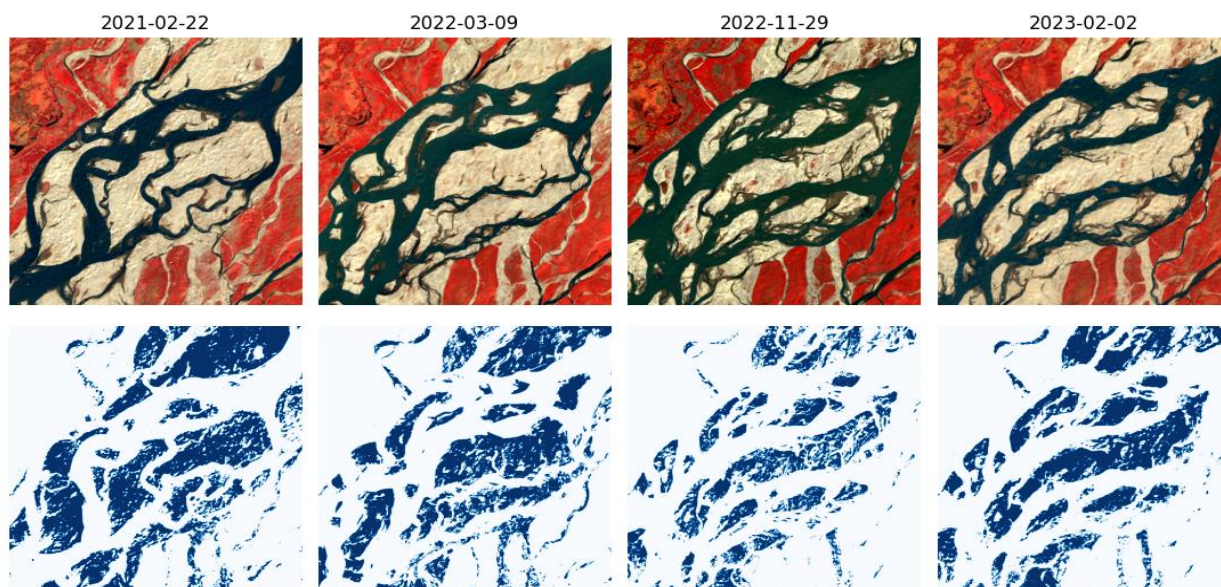


Рисунок 6. – Візуалізація результатів розпізнавання води на супутникових знімках за допомогою нечіткої нейронної мережі Такагі-Сугено (TS-FNN)

Нечіткі правила в TS-FNN краще інтерпретувати, оскільки вихідні дані є лінійною функцією вхідних змінних, що полегшує розуміння того, як модель приймає свої рішення.

3.2.4. Модель нечіткої нейронної мережі NEFCLASS

NEFCLASS – це нейро-нечітка система, призначена для завдань класифікації. Він поєднує нечітку логіку зі структурою прямої нейронної мережі, де нечіткі правила вивчаються на основі вхідних даних. Модель автоматично генерує нечіткі правила класифікації, які потім налаштовуються за допомогою методів навчання нейронної мережі. нейронна мережа коригує нечіткі правила, щоб мінімізувати помилку класифікації.

Для простоти ми реалізували функціональність, подібну до NEFCLASS, за допомогою нейронної мережі для представлення нечітких правил, але ми вручну визначили процес фазифікації. Результат вказаний на рисунку 7.

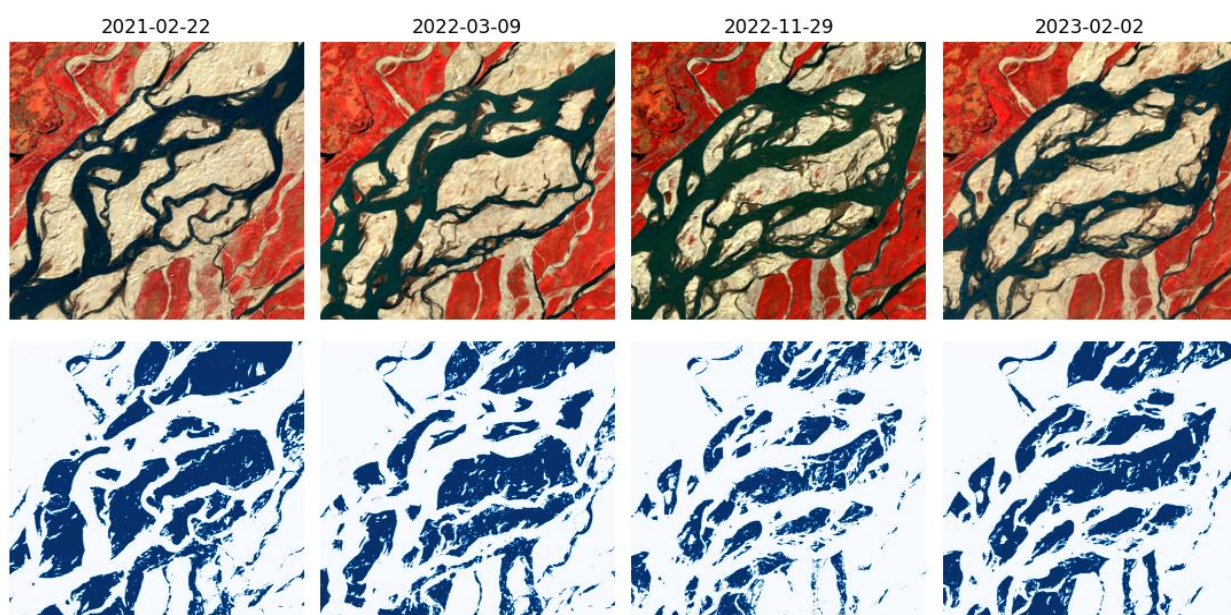


Рисунок 7. – Візуалізація результатів розпізнавання води на супутникових знімках за допомогою нечіткої нейронної мережі NEFCLASS

Для реалізації даної моделі ми використали той же метод, що і у попередніх моделях, де визначили нечіткі набори для червоного та інфрачервоного діапазонів (низький, середній, високий). Нейронна мережа побудована так, щоб вивчати нечіткі правила неявно. Вхідний рівень відповідає нечітким входам (нечітким характеристикам), а наступні шари моделюють комбінацію нечітких правил.

Модель навчається з використанням `categorical_crossentropy` функції втрат, оскільки це завдання класифікації, а кінцевим результатом є функція активації `softmax`, яка виводить ймовірність кожного кластера для заданого вхідного сигналу. Після навчання мережа прогнозує найбільш імовірний кластер для кожного пікселя в повному наборі даних. Прогнозовані призначення кластерів змінюються та візуалізуються. Прогнозовані кластери візуалізуються так само, як і раніше, де кожному кластеру призначається певний колір, і кластери відображаються для кожного супутникового зображення.

Найбільшими перевагами даної моделі є можливість вивчати нечіткі правила під час навчання, що робить її придатною для обробки складних рішень, типових для класифікації супутникових зображень. NEFCLASS може масштабуватися до більших наборів даних завдяки структурі нейронної мережі, яка дозволяє йому ефективніше обробляти багатовимірні дані.

3.2.5. Гібридна модель NEFCLASS- EM

З метою досягнення найвищої точності кластеризації, ми також перевірили гібридизацію мережі NEFCLASS алгоритмом електромагнітної метаевристики. Результат вказаний на рисунку 8.

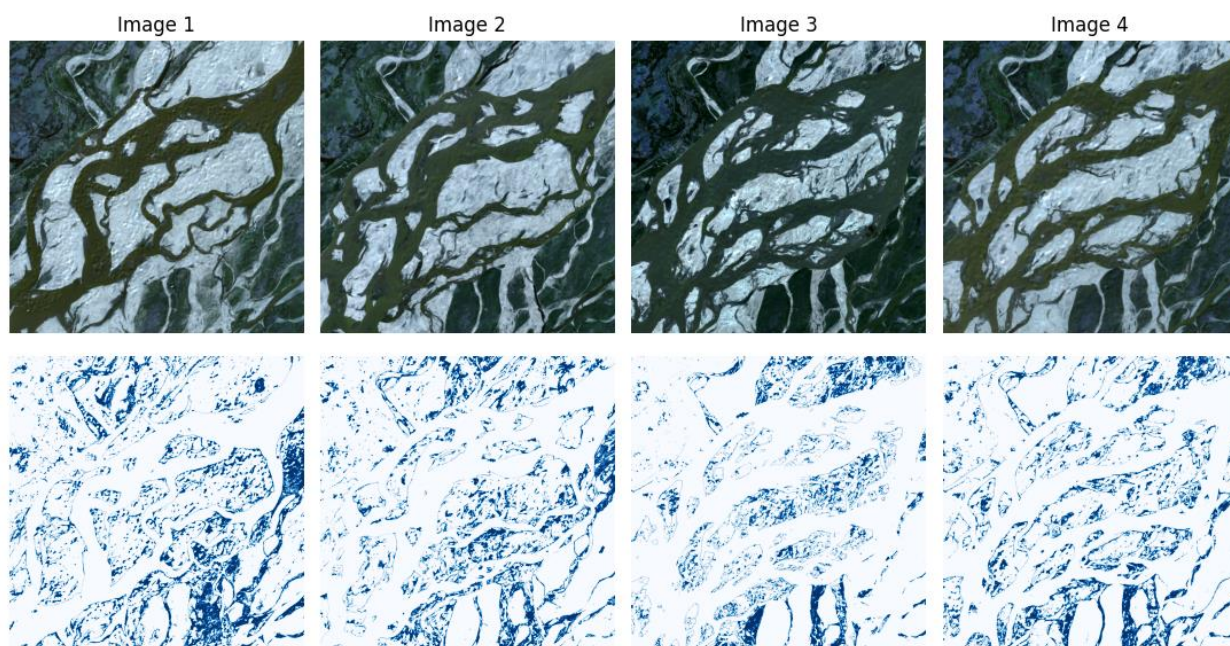


Рисунок 8. – Візуалізація результатів розпізнавання води на супутникових знімках за допомогою гібридної нейронної мережі NEFCLASS- EM

Щоб гібридувати задану нейронну мережу алгоритмом електромагнітної метаевристики (ЕМ), ми розробили стратегію оптимізації на основі пошуку з процесом навчання нейронної мережі. Зокрема, алгоритм ЕМ застосували для оптимізації гіперпараметрів моделі NEFCLASS, покращуючи ефективність класифікації. Алгоритм ЕМ моделює силу притягання та відштовхування між кандидатськими рішеннями (ваговими конфігураціями) на основі електромагнітних принципів. Оптимізовані вагові коефіцієнти

використали для подальшого навчання мережі NEFCLASS за допомогою оптимізатора Адама. Дана гібридизація призвела до швидшої конвергенції та кращої точності класифікації за рахунок використання можливостей глобального пошуку ЕМ та можливості тонкого налаштування оптимізації на основі градієнта.

3.3. Порівняння отриманих результатів та оцінка точності моделей нечітких нейронних мереж

Нечітка мережа FNN демонструє значне покращення точності, швидко перевищивши 91% до другої епохи та стабілізувавшись приблизно на 96% до 20-ї епохи. Втрати також стабільно зменшуються, що вказує на ефективне навчання та зменшення помилок. TS-FNN демонструє швидке зменшення втрат, досягаючи 0,0114 до 20-ї епохи. Ця швидка конвергенція свідчить про високу точність. NEFCLASS показує результат, подібний до FNN, але починає з дещо меншою точністю. До 20-ї епохи вона досягає точності понад 95%. Зменшення втрат відбувається повільніше порівняно з FNN, але все ще суттєве. NEFCLASS-EU починає з найнижчої точності та найвищих втрат, але швидко стабілізується. До кінця навчання модель досягає продуктивності, подібної до NEFCLASS, з точністю понад 95% і значним зменшенням втрат. Результати точності навчання моделей вказані в таблиці 1, результати втрат під час навчання вказані в таблиці 2.

Таблиця 1. – Порівняння отриманих результатів точності навчання нечітких нейронних мереж

№ Епохи	Метод			
	Нечітка нейронна мережа (FNN)	Нечітка нейронна мережа Такагі- Сугено (TS-FNN)	Нечітка нейронна мережа NEFCLASS	Гібридна NEFCLASS- EM
	Точність навчання FNN			
1.	0.5781	0.5710	0.5630	0.4789
2.	0.9162	0.9113	0.8616	0.8452
3.	0.9488	0.9363	0.9102	0.9068
4.	0.9578	0.9368	0.9177	0.9237
5.	0.9594	0.9391	0.9227	0.9289
6.	0.9582	0.9382	0.9279	0.9355
7.	0.9584	0.9404	0.9330	0.9425
8.	0.9647	0.9563	0.9398	0.9463
9.	0.9630	0.9590	0.9369	0.9464
10.	0.9662	0.9562	0.9466	0.9478
11.	0.9642	0.9558	0.9437	0.9498
12.	0.9639	0.9539	0.9445	0.9505
13.	0.9639	0.9539	0.9513	0.9553
14.	0.9640	0.9540	0.9444	0.9531
15.	0.9626	0.9526	0.9506	0.9547
16.	0.9643	0.9543	0.9488	0.9587
17.	0.9643	0.9543	0.9485	0.9569
18.	0.9664	0.9564	0.9452	0.9544
19.	0.9619	0.9519	0.9543	0.9534
20.	0.9679	0.9579	0.9512	0.9560

Таблиця 2. – Порівняння отриманих результатів втрат під час навчання нечітких нейронних мереж

№ Епохи	Метод			
	Нечітка нейронна мережа (FNN)	Нечітка нейронна мережа Такагі-Сугено (TS-FNN)	Нечітка нейронна мережа NEFCLASS	Гібридна NEFCLASS- EM
	<i>Втрати під час навчання FNN</i>			
1.	1.1960	0.1200	1.3084	1.3472
2.	0.2958	0.0320	0.4620	0.4864
3.	0.1422	0.0230	0.2614	0.2674
4.	0.1114	0.0199	0.2085	0.2167
5.	0.1052	0.0175	0.1849	0.1815
6.	0.1039	0.0160	0.1772	0.1658
7.	0.0945	0.0145	0.1640	0.1543
8.	0.0889	0.0145	0.1521	0.1389
9.	0.0893	0.0138	0.1482	0.1357
10.	0.0848	0.0141	0.1350	0.1296
11.	0.0834	0.0139	0.1371	0.1256
12.	0.0848	0.0124	0.1403	0.1256
13.	0.0827	0.0126	0.1255	0.1167
14.	0.0857	0.0126	0.1299	0.1259
15.	0.0850	0.0124	0.1189	0.1140
16.	0.0823	0.0120	0.1157	0.1093
17.	0.0846	0.0122	0.1199	0.1062
18.	0.0796	0.0115	0.1195	0.1064
19.	0.0848	0.0113	0.1099	0.1075
20.	0.0771	0.0114	0.1141	0.1073

Отже, з отриманих результатів навчання нечітких нейронних мереж ми бачимо, що FNN демонструє найкращий загальний баланс швидкого підвищення точності та стабільного зменшення втрат, що робить його найефективнішою моделлю з точки зору як швидкості навчання, так і кінцевої продуктивності. TS-FNN має найшвидше зменшення втрат, що вказує на дуже точну модель. NEFCLASS можна порівняти з FNN, але трохи повільніше з точки зору зростання точності та зменшення втрат. NEFCLASS- EM починає з найнижчої продуктивності, але наздоганяє, досягаючи рівнів точності, близьких до FNN і NEFCLASS до останньої епохи.

Підсумовуючи, FNN і TS-FNN видаються найсильнішими моделями, а NEFCLASS і NEFCLASS- EM трохи відстають, але все ще демонструють сильну конвергенцію.

3.3. Оцінка якості води та потенціалу підземних вод

В результаті проведених досліджень ми запропонували методику оцінки якості води за допомогою супутникових зображень та методів обчислювального інтелекту, а саме нечітких нейронних мереж. Даний метод передбачає поєднання методів дистанційного зондування та геоінформаційного моделювання з передовими моделями машинного навчання та штучного інтелекту (ШІ). Цей підхід дозволяє проводити широкомасштабний автоматизований аналіз параметрів якості води, таких як каламутність, концентрація хлорофілу-а та загальна кількість завислих речовин (TSS).

Для даного методу ми визначили такі чотири діапазони якості води:

- 1) Синій (450–500 nm): чутливий до хлорофілу-а та прозорості води.
- 2) Зелений (500–600 nm): відображає органічні речовини та зважені частинки.
- 3) Червоний (600–700 nm): допомагає виявити відкладення.
- 4) Близький інфрачервоний діапазон (700–1100 nm): корисний для визначення каламутності та TSS.

Наступним кроком ми вилучили спектральні та просторові характеристики із супутникових зображень, які мають значення для оцінки якості води. Дані характеристики є вхідними даними для моделей обчислювального інтелекту. Оцінку хлорофілу ми отримали з обчисленням нормалізованого різницевого індексу хлорофілу (NDCI) (формула 16):

$$NDCI = \frac{Red - Green}{Red + Green} \quad (16)$$

Для оцінки каламутності ми використали індекс каламутності, отриманий з червоного та ближнього інфрачервоного діапазону. Поєднання червоних і зелених смуг допомогло оцінити загальну кількість завислих речовин (TSS). Для покращення аналізу шляхом фіксації просторових змін у водоймах, ми використали функції просторової текстури (матрицю суміжних рівнів сірого). Для кращого розуміння характеристик водойм – розраховували статистичні характеристики, такі як середнє значення, дисперсію та ентропію значень пікселів.

Для калібрування та перевірки моделей обчислювального інтелекту використали дані наземних вимірювань параметрів якості води, таких як хлорофіл, каламутність та кількість завислих речовин.

Ідея з застосуванням нечітких нейронних мереж принесла кращі результати, ніж результати, які базуються на традиційних нейронних мережах та традиційних методах машинного навчання. Саме системи нечіткої логіки ідеально підходять для обробки невизначеності в оцінці якості води, особливо коли межі між класами якості нечіткі. Нечіткі нейронні мережі дають можливість визначати нечіткі функції приналежності для параметрів якості води на основі супутникових функцій (наприклад, піксель може належати як до «чистої води», так і до «забрудненої води» з різними ступенями приналежності).

Після навчання та перевірки моделі її можна використовувати для прогнозування параметрів якості води за новими супутниковими знімками. Та створювати карти якості води. Застосовуючи моделі до супутникових зображень за різні періоди часу, можна відстежувати тимчасові зміни якості води. Це особливо корисно для виявлення таких тенденцій, як цвітіння водоростей, забруднення або сезонні коливання якості води. Безперервно контролюючи якість води, моделі обчислювального інтелекту можуть допомогти виявити ранні ознаки забруднення або цвітіння водоростей і викликати попередження.

Отримання правильних характеристик із супутникових зображень має вирішальне значення для оцінки потенціалу підземних вод. Дані функції часто служать вхідними даними для моделей обчислювального інтелекту. Наприклад, Нормалізований індекс різниці рослинності (NDVI) допомагає оцінити здоров'я та щільність рослинності, яка пов'язана з наявністю ґрунтових вод.

$$NDVI = \frac{NIR - Red}{NIR + Red} \quad (17)$$

Цифрові моделі рельєфу (Рис. 9(а)) допомагають зрозуміти рельєф місцевості так як низинні території та долини мають вищий потенціал поповнення підземних вод. Нахил впливає на стік води, а висота (напрямок схилу) впливає на утримання вологи. Аналіз річкових моделей і моделей дренажу за допомогою супутникових даних допомагає визначити території з високим потенціалом інфільтрації.

Поверхнева вологість ґрунту (SSM): отримана за допомогою таких супутників, як Sentinel-1 (Рис. 9(б)), показує, скільки води утримує ґрунт. Висока вологість ґрунту передбачає хороші умови поповнення ґрунтових вод. Вміст води в рослинності може опосередковано вказувати на рівень вологості ґрунту та потенційну наявність ґрунтових вод.

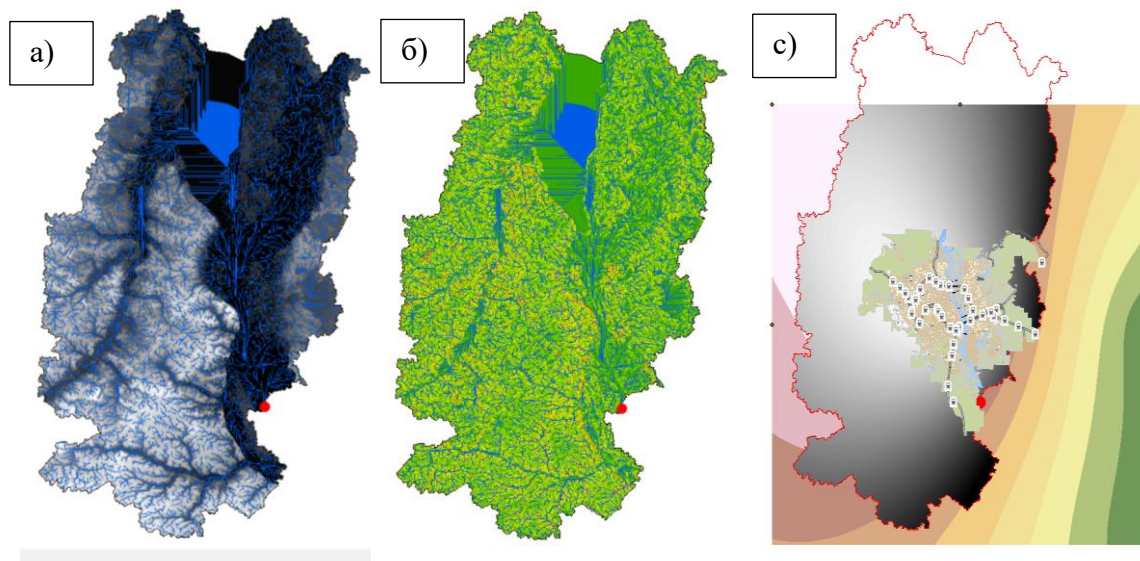


Рисунок 9. – Візуалізація результатів отримання цифрової моделі рельєфу (а), вологості ґрунту (б),

Супутникові радарні та оптичні зображення можуть допомогти ідентифікувати геологічні лінеamenti (розломи, тріщини), які діють як канали для ґрунтових вод. Для цієї мети можуть бути корисні радіолокаційні дані супутників, таких як Sentinel-1. Визначення різних типів гірських порід або утворень (наприклад, пористих порід, тріщинуватих водоносних горизонтів) за супутниковими даними є критично важливим, оскільки певні геологічні утворення є більш сприятливими для зберігання підземних вод. Картування озер, річок і водно-болотних угідь із супутникових зображень дає змогу зрозуміти зони поповнення підземних вод (рис. 10), оскільки поверхневі водойми часто пов'язані з водоносними горизонтами.

Для навчання та перевірки моделей обчислювального інтелекту ми зібрали наземні дані, такі як: рівні ґрунтових вод з свердловин, профілі вологості ґрунту, геологічні та гідрогеологічні дослідження та основні кліматичні дані (кількість опадів, швидкість випаровування), щоб визначити потенціал поповнення.

Після виділення характеристик з супутникових знімків ми застосували методи обчислювального аналізу для моделювання зв'язків між цими характеристиками та потенціалом підземних вод. Такі характеристики як: індекс NDVI, DEM, вологість ґрунту, землекористування, нахил та щільність дренажу використовувались як вхідні змінні. В результаті отримали модель, яка здатна прогнозувати території з високим, або низьким потенціалом ґрунтових вод на основі вхідних характеристик.

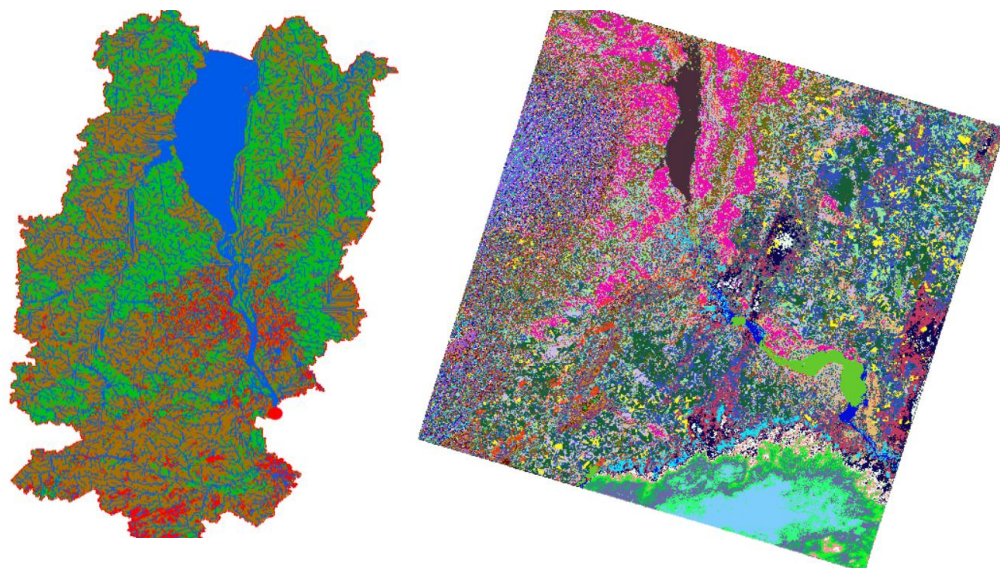


Рисунок 10. – Результат картування озер, річок і водно-болотних угідь із супутникових зображень

Нечітка логіка особливо корисна в оцінці ґрунтових вод, оскільки вона може обробляти невизначеність і мінливість даних (наприклад, змінна вологість ґрунту). Просторова та часова роздільна здатність супутникових даних може вплинути на точність оцінки ґрунтових вод. Дані високої роздільної здатності підвищують точність, але можуть бути доступними не в усіх регіонах.

Оцінка потенціалу підземних вод за допомогою супутникових зображень і методів обчислювального аналізу поєднує в собі сильні сторони дистанційного зондування, машинного навчання та науки про дані. Завдяки об'єднанню ключових характеристик, таких як вологість ґрунту, показники рослинності, рельєф і геологічні структури, ці методи забезпечують потужний, економічно ефективний підхід до дослідження та моніторингу підземних вод.

4. ВИСНОВКИ

Оцінка якості води за супутниковими знімками за допомогою методів обчислювального інтелекту є потужним підходом для широкомасштабного автоматизованого моніторингу водойм. Поєднання даних дистанційного зондування з такими методами, як штучні нейронні мережі, нечітка логіка, дерева рішень і опорні векторні машини, дозволяє точно прогнозувати ключові параметри якості води, такі як хлорофіл-а, каламутність і загальний вміст завислих речовин. Завдяки безперервному моніторингу якості води за допомогою супутникових зображень методи обчислювального інтелекту можуть допомогти виявити зміни навколишнього середовища, керувати водними ресурсами та запобігати забрудненню води.

В результаті проведеного дослідження можна зазначити такі ключові моменти та переваги запропонованих в статті методів:

- дистанційне зондування дозволяє безперервно охоплювати широку територію водойм і наземних ландшафтів, уможливорюючи широкомасштабний моніторинг, який був би непрактичним за допомогою традиційних наземних методів. Супутникові дані дають важливу інформацію про якість води та потенціал підземних вод у величезних регіонах, включаючи віддалені та важкодоступні райони.

➤ методи обчислювального інтелекту, такі як нечіткі нейронні мережі, можуть автоматизувати процес аналізу складних супутникових даних, значно скорочуючи час і робочу силу, необхідні для оцінки водних ресурсів. Ці моделі можуть швидко виявляти закономірності, аномалії та тенденції, надаючи інформацію в режимі реального або майже в реальному часі, яка є критичною для своєчасного прийняття рішень.

➤ нечіткі нейронні мережі чудово справляються з невизначеністю та неточністю, які притаманні природному середовищу. Параметри якості води та потенціал підземних вод часто демонструють складні, нелінійні зв'язки, на які впливають численні фактори (наприклад, землекористування, рослинність, клімат). FNN ефективно моделюють ці складності та забезпечують м'які класифікації, які дозволяють робити гнучкіші прогнози

➤ використання супутникових даних і моделей на основі штучного інтелекту зводить до мінімуму потребу в розширених польових дослідженнях, зменшуючи витрати, пов'язані з традиційною оцінкою якості води та ґрунтових вод. Завдяки використанню вільно доступних супутникових платформ (наприклад, Sentinel, Landsat) і потужних обчислювальних засобів моніторинг навколишнього середовища стає доступнішим і масштабованим.

➤ геоінформаційне моделювання дозволяє інтегрувати просторові та часові дані, дозволяючи відслідковувати тенденції якості води та потенціал підземних вод з часом. Це допомагає визначити сезонні закономірності, довгострокові зміни навколишнього середовища та наслідки людської діяльності, наприклад сільського господарства, урбанізації чи забруднення.

Отже, поєднання нечітких нейронних мереж, дистанційного зондування Землі та геоінформаційного моделювання пропонує надійну, динамічну та ефективну структуру для управління та охорони водних ресурсів, забезпечуючи їх сталість перед зростанням екологічних проблем.

ПОДЯКИ

Ідеї для даного дослідження базуються на знаннях з оцінки якості води, поглиблених та розширених під час проходження курсу “Operation and Maintenance of Urban Water Supply System (Water Quality and Purification) (B)” в Японії, з фінансовою підтримкою Японського агентства міжнародного співробітництва (JICA).

СПИСОК ВИКОРИСТАНИХ ДЖЕРЕЛ

1. [Jinyue Chen, Shuisen Chen, Rao Fu, Dan Li, Hao Jiang, Chongyang Wang, Yongshi Peng, Kai Jia, Brendan J. Hicks](#). Remote Sensing Big Data for Water Environment Monitoring: Current Status, Challenges, and Future Prospects First published: 21 January 2022 <https://doi.org/10.1029/2021EF002289>

2. Hong, H. Y., Jaafari, A., & Zenner, E. K. (2019). Predicting spatial patterns of wildfire susceptibility in the Huichang County, China: An integrated model to analysis of landscape indicators. *Ecological Indicators*, **101**, 878–891. <https://doi.org/10.1016/j.ecolind.2019.01.056>

3. Tang, L. L., Zhang, S., Zhang, J. H., Liu, Y., & Bai, Y. (2021). Estimating evapotranspiration based on the satellite-retrieved near-infrared reflectance of vegetation (NIRv) over croplands. *GIScience and Remote Sensing*, **58**(6), 889–913. <https://doi.org/10.1080/15481603.2021.194762>

4. UN General Assembly. Transforming Our World: The 2030 Agenda for Sustainable Development. 21 October 2015. Available online: <https://www.refworld.org/docid/57b6e3e44.html> (accessed on 5 December 2022).

5. European Parliament. Directive 2000/60/EC—Framework for Community Action in the Field of Water Policy; European Parliament: Bruxelles, Belgium, 2003. [[Google Scholar](#)]
6. He, J.; Chen, Y.; Wu, J.; Stow, D.A.; Christakos, G. Space-Time Chlorophyll-a Retrieval in Optically Complex Waters that Accounts for Remote Sensing and Modeling Uncertainties and Improves Remote Estimation Accuracy. *Water Res.* **2019**, *171*, 115403. [[Google Scholar](#)] [[CrossRef](#)]
7. Nas, B.; Ekercin, S.; Karabörk, H.; Berktaş, A.; Mulla, D.J. An Application of Landsat-5TM Image Data for Water Quality Mapping in Lake Beyşehir, Turkey. *Water Air Soil Pollut.* **2010**, *212*, 183–197. [[Google Scholar](#)] [[CrossRef](#)]
8. Govedarica, M.; Jakovljevic, G. Monitoring spatial and temporal variation of water quality parameters using time series of open multispectral data. In Proceedings of the SPIE 11174 Seventh International Conference on Remote Sensing and Geoinformation of the Environment, Paphos, Cyprus, 18–21 March 2019. [[Google Scholar](#)]
9. Wu, C.; Wu, J.; Qi, J.; Zhang, L.; Huang, H.; Lou, L.; Chen, Y. Empirical estimation of total phosphorus concentration in the mainstream of the Qiantang River in China using Landsat TM data. *Int. J. Remote Sens.* **2010**, *31*, 2309–2324. [[Google Scholar](#)] [[CrossRef](#)]
10. Gordana Jakovljevic, Flor Álvarez-Taboada, Miro Govedarica. Long-Term Monitoring of Inland Water Quality Parameters Using Landsat Time-Series and Back-Propagated ANN: Assessment and Usability in a Real-Case Scenario. *Remote Sens.* **2024**, *16*(1), 68; <https://doi.org/10.3390/rs16010068>
11. Ha, N.; Koike, K.; Nhuan, M. Improved Accuracy of Chlorophyll-a Concentration Estimates from MODIS Imagery Using a Two-Band Ratio Algorithm and Geostatistics: As Applied to the Monitoring of Eutrophication Processes over Tien Yen Bay (Norther Vietnam). *Remote Sens.* **2013**, *6*, 421–442. [[Google Scholar](#)] [[CrossRef](#)]
12. Nechad, B.; Ruddick, K.; Park, Y. Calibration and validation of a generic multisensor algorithm for mapping of total suspended matter in turbid waters. *Remote Sens. Environ.* **2010**, *114*, 854–866. [[Google Scholar](#)] [[CrossRef](#)]
13. Brezonik, P.L.; Olmanson, L.G.; Finlay, J.C.; Bauer, M.E. Factors Affecting the Measurement of CDOM by Remote Sensing of Optically Complex Inland Waters. *Remote Sens. Environ.* **2015**, *157*, 199–215. [[Google Scholar](#)] [[CrossRef](#)]
14. Shahzad, M.I.; Meraj, M.; Nazeer, M.; Zia, I.; Inam, A.; Mehmood, K.; Zafar, H. Empirical Estimation of Suspended Solids Concentration in the Indus Delta Region Using Landsat-7 ETM+ Imagery. *J. Environ. Manag.* **2018**, *209*, 254–261. [[Google Scholar](#)] [[CrossRef](#)]
15. Guo, H.; Huang, J.J.; Zhu, X.; Wang, B.; Tian, S.; Xu, W.; Mai, Y. A generalized machine learning approach for dissolved oxygen estimation at multiple spatiotemporal scales using remote sensing. *Environ. Pollut.* **2021**, *288*, 117734. [[Google Scholar](#)] [[CrossRef](#)] [[PubMed](#)]
16. Qian, J.; Liu, H.; Qian, L.; Bauer, J.; Xue, X.; Yu, G.; He, Q.; Zhou, Q.; Bi, Y.; Norra, S. Water quality monitoring and assessment based on cruise monitoring, remote sensing, and deep learning: A case study of Qingcaosha Reservoir. *Front. Environ. Sci.* **2022**, *10*, 979133. [[Google Scholar](#)] [[CrossRef](#)]
17. Guo, H.; Huang, J.J.; Chen, B.; Guo, X.; Singh, V.P. A machine learning-based strategy for estimating non-optically active water quality parameters using Sentinel-2 imagery. *Int. J. Remote Sens.* **2020**, *42*, 1841–1866. [[Google Scholar](#)] [[CrossRef](#)]
18. Peterson, K.T.; Sagan, V.; Sloan, J.J. Deep learning-based water quality estimation and anomaly detection using Land-sat-8/Sentinel-2 virtual constellation and cloud computing. *Giscience Remote Sens.* **2020**, *57*, 510–525. [[Google Scholar](#)] [[CrossRef](#)]
19. Hafeez, S.; Wong, M.S.; Ho, H.C.; Nazeer, M.; Nichol, J.E.; Abbas, S.; Tang, D.; Lee, K.-H.; Pun, L. Comparison of Machine Learning Algorithms for Retrieval of Water Quality Indicators in Case-II Waters: A Case Study of Hong Kong. *Remote Sens.* **2019**, *11*, 617. [[Google Scholar](#)] [[CrossRef](#)]

20. Peterson, K.T.; Sagan, V.; Sloan, J.J. Deep learning-based water quality estimation and anomaly detection using Land-sat-8/Sentinel-2 virtual constellation and cloud computing. *Giscience Remote Sens.* **2020**, *57*, 510–525. [[Google Scholar](#)] [[CrossRef](#)]
21. Gómez, D.; Salvador, P.; Sanz, J.; Casanova, J.L. A new approach to monitor water quality in the Menor sea (Spain) using satellite data and machine learning methods. *Environ. Pollut.* **2021**, *286*, 117489. [[Google Scholar](#)] [[CrossRef](#)] [[PubMed](#)]
22. Balabin, R.M.; Lomakina, E.I. Support vector machine regression (SVR/LS-SVM)—An alternative to neural networks (ANN) for analytical chemistry? Comparison of nonlinear methods on near infrared (NIR) spectroscopy data. *Analyst* **2011**, *136*, 1703–1710. [[Google Scholar](#)] [[CrossRef](#)] [[PubMed](#)]
23. Jakovljevic, G.; Govedarica, M.; Alvarez-Taboada, F. Water body mapping: A comparison of remotely sensed and GIS open data sources. *Int. J. Remote Sens.* **2018**, *40*, 2936–2964. [[Google Scholar](#)] [[CrossRef](#)]
24. USGS. Landsat 4-7 Collection 1 Surface Reflectance Code LEDAPS Product Guide. Available online: https://d9-wret.s3.us-west-2.amazonaws.com/assets/palladium/production/s3fs-public/atoms/files/LSDS-1370_L4-7_C1-SurfaceReflectance-LEDAPS_ProductGuide-v3.pdf (accessed on 25 November 2022).
25. Wang, X. *MATLAB Neural Networks 43 Case Studies*, 1st ed.; Beijing University of Aeronautics and Astronautics Press: Beijing, China, 2013; pp. 288–289. [[Google Scholar](#)]
26. Zhou, Z. The Application of Fuzzy Neural Network Based on T-S Model in Water Quality Evaluation. Master's Thesis, East China University of Political Science and Law, Nanjing, China, 2007. [[Google Scholar](#)]