

УДК 681.58

Путренко Віктор ВалентиновичКандидат географічних наук, докторант, orcid.org/0000-0002-0239-9241

Національний технічний університет України «Київський політехнічний інститут ім. Ігоря Сікорського», Київ

Пашинська Наталія МиколаївнаКандидат географічних наук, старший науковий співробітник кафедри інтелектуальних та інформаційних систем, orcid.org/0000-0002-0133-688X

Київський національний університет ім. Тараса Шевченка, Київ

**ІНТЕЛЕКТУАЛЬНИЙ АНАЛІЗ ЗЕМНОГО ПОКРИВУ ТЕРИТОРІЇ УКРАЇНИ
НА ОСНОВІ ДАНИХ GLOBELAND 30**

***Анотація.** В роботі розглянуто аналіз та типізацію земного покриву території України для цілей державного територіального управління та адміністрування з використанням даних продукту глобальної класифікації земного покриву Globeland 30. Для проведення типізації використовувалися набір методів ГІС-аналізу, зокрема обробки растрових наборів даних, отримання тайлових мозаїк та конвертації, обробки векторних даних шляхом перекласифікації та зональної статистики, методи кластеризації, зокрема метод побудови карт самоорганізації Кохонена для цілей кластеризації та типізації адміністративних одиниць. На основі цих даних проведено типізацію адміністративних одиниць та отримано розподіл на 9 кластерів, які відображають основні типи землекористування в Україні. Таким чином, було отримано групи районів, які за своїм територіальним положенням, спеціалізацією та особливостями розвитку мають подібні структури землекористування. Отримані результати типізації є основою для формування та корегування програм територіального розвитку на рівні адміністративних одиниць, контролю за природоохоронною, сільськогосподарською та містобудівною діяльністю.*

***Ключові слова:** інтелектуальний аналіз; супутникові знімки; земний покрив; землекористування; ГІС-аналіз; карти самоорганізації Кохонена; кластеризація*

Вступ

Отримання актуальної та достовірної геопросторової інформації про особливості території є запорукою проведення успішного планування територіального розвитку на переході до сталих форм господарювання. В цьому контексті забезпечення органів державного управління та планування якісними геопросторовими даними є одним із найважливіших завдань. Джерелами таких даних у більшості випадків виступають дані дистанційного зондування Землі (ДЗЗ), які після певної технічної підготовки можуть використовуватися для вирішення прикладних завдань моніторингу та аналізу території. В останні роки все більшого значення набувають дані космічної зйомки, а також широко використовуються результати лідарної зйомки та БПЛА різного класу для отримання знімків у режимі близькому до реального часу.

Одним із базових продуктів, який може бути отримано на основі обробки даних ДЗЗ є карта земного покриву або землекористування. Подібні дані становлять основу для організації та управління територією. Особливо важливо мати динамічний ряд

для відслідковування тенденцій в зміні видів землекористування. Враховуючи важливість цих даних, останні 30 років у всьому світі реалізуються міжнародні та регіональні програми та ініціативи, що направлені на дослідження земного покриву та створення карт землекористування. Найбільш відомим та в певному змісті піонерським стало отримання світового покриття GLC2000 (рис. 1) та успішна реалізація європейської програми CORINE, яка заклала єдині стандарти відображення типів землекористування в Європейському Союзі. Оскільки Україна не брала участь у програмі CORINE територія країни довгий час залишалась без якісного опису земного покриву. Відповідно до цього територіальне планування та регіональне управління залишались без інформації про землекористування і відповідно поступались європейським методам прийняття рішень.

Ситуація змінилась зі створенням у Китаї глобального земного покриття високого розрізнення – Globeland 30. Первинні дані покриття доступні для використання та аналізу з метою оцінки земного покриву. Враховуючи, що дана модель отримання даних землекористування була відпрацьована на основі наземних даних в Китаї та недостатньо

враховує українські географічні особливості, ці дані можуть містити певні похибки класифікації, але все одно можуть бути з високою ефективністю використані для проведення регіонального аналізу землекористування (рис. 1).

Аналіз публікацій

За останні роки з'явився цілий ряд робіт, які присвячені аналізу та використанню продуктів класифікації земного покриву. Значна кількість європейських робіт зосереджена на аналізі покриття CORINE, яке є базовим для управління землекористуванням в Європейському Союзі [3; 5; 10]. Основна увага в публікаціях приділена питанням алгоритмів класифікації земного покриву та методам верифікації даних [4; 8]. Подібні роботи проводились для території України з метою покращення якості та точності даних. На основі використання цих даних створено ряд нових методів обробки даних із застосуванням методів machine learning [1; 2]. Останні розробки спрямовані на залучення краудсорсингового збору даних про земний покрив з метою верифікації та навчання систем класифікації [7]. При цьому відкритим залишається питання подальшого застосування даних класифікації для розробки регіональної політики на засадах концепції сталого розвитку.

Мета статті

Метою дослідження є аналіз земного покриву території України для цілей державного управління та адміністрування з використанням даних продукту Globeland 30. Завданнями дослідження є вивчення особливостей продукту Globeland 30, методів отримання даних земного покриву, розробка схеми та методики обробки даних для задач дослідження території України, отримання та інтелектуальний аналіз зональних даних для адміністративних одиниць України з використанням карт самоорганізації Кахонена.

Виклад основного матеріалу

З метою підтримки досліджень глобальних змін і розвитку системних моделей функціонування Землі, Міністерство науки і технологій Китаю запустило Національну високотехнологічну програму досліджень та розробок «Високі технології досліджень і розробки Програми Китаю» (863 Program) під назвою «Картографування за допомогою даних ДЗЗ і дослідження ключових технологій побудови глобального рослинного покриву» у 2010 році. Одним з результатів проекту є картографічний продукт GlobeLand30-2010 глобального покриву з 30-метровим просторовим розрізненням, отриманий із зображень дистанційного зондування в 2010 році [9].

Набір даних охоплює територію площею від 80° Пн. ш. до 80° Пд. ш. та складається з 10 типів земного покриву, а саме сільськогосподарські землі, лісові, степові, деревно-чагарникові, водно-болотні угіддя, водні об'єкти, тундра, урбаністичні ландшафти, пустельні ландшафти та льодовики.

Для генерації даних земного покриття використовувалися класифіковані зображення зі супутникових даних мультиспектральних зображень GlobeLand30-2010 з роздільною здатністю 30 м, в тому числі Landsat TM і ETM+ і мультиспектральні зображення китайської системи Environmental Disaster Alleviation Satellite (HJ-1). Були обрані безхмарні зображення, отримані протягом вегетаційного сезону в межах ± 1 рік відносно 2010 року.

Усі зображення Landsat TM і ETM+ були завантажені з ресурсу Геологічної служби США (USGS, <http://landsat.usgs.gov/>) на рівні 1T. До знімків Landsat були застосовані геометричні і радіометричні виправлення зображень, які потім були зареєстровані в системі координат WGS-84 системи з проекцією UTM. Таким чином, були завантажені 9907 зображення Landsat TM і ETM+ зображень з USGS, близько 80% з них були придбані в період з 2009 по 2011 р.

Попередня обробка Landsat TM і ETM+ зображень включає в себе атмосферну і топографічну корекції. Атмосферна корекція використовує оновлену версію алгоритму Швидкої лінії візування атмосферного аналізу спектрального Гіперкубу (FLAASH), а топографічна корекція використовує метод згладжених COS (SCOS).

HJ-1 китайський супутник дистанційного зондування, спрямований на моніторинг навколишнього середовища та запобігання і зменшення стихійних лих. HJ-1A супутник несе мультиспектральну камеру (CCD) і гіперспектральний сенсор (HSI). Супутник HJ-1B несе мультиспектральну камеру (CCD) та інфрачервоний мультиспектральний сенсор (IRS) (<http://www.cresda.com/>).

1465 сцен мультиспектральних зображень HJ-1 було використано для отримання інформації про земний покрив, покриваючи 60% від загальної площі земної кулі, з часом придбання, починаючи з вересня 2008 року по грудень 2011 року.

Мінімальна одиниця відображення (MOB) належить до найменшої ділянки, яка використовується для контролю якості. Площа ділянки більша за розмір MOB повинна пройти через процес контролю якості. Що стосується лінійних об'єктів, таких як річки, для контролю застосовується мінімальна ширина. MOB варіюється від типу до типу земного покриву, який визначається такими факторами:

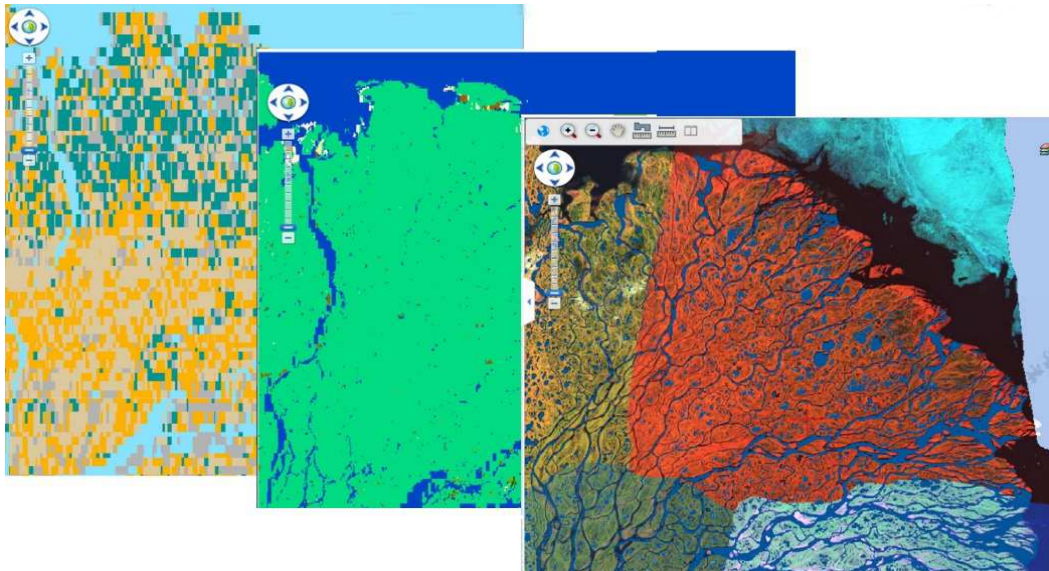


Рисунок 1 – Порівняння даних глобальних продуктів земного покриття GLC2000, GLC2009, GlobeLand30

- просторовий розподіл об'єктів (розмір, щільність і т.д.);
- масштабні характеристики типу об'єктів;
- обсяг роботи людино-машинного інтерактивного редагування та перевірки;
- мінімальні вимоги до точності класифікації об'єктів.

Продукт GlobeLand30-2010 було отримано з використанням ієрархічного методу вилучення.

Кожен тип рослинного покриття було класифіковано по одному, з обмеженнями маскою інших типів рослинного покриття. Робочий процес полягає в тому щоб отримати тільки один тип земного покриття за один раз, а потім маскувати клас після детекції. Далі класифікація здійснюється для наступного типу рослинного покриття, який потім маскується, допоки не будуть отримані всі класи (рис. 2).

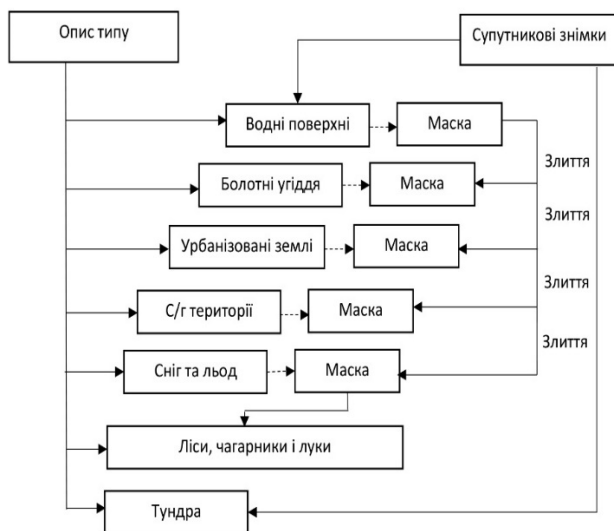


Рисунок 2 – Ієрархічна стратегія класифікації типів

Порядок отримання типів земного покриття такий:

1. Водойми;
2. Водно-болотні угіддя;
3. Урбанізовані території;
4. Сільськогосподарські землі;
5. Постійні сніги та лід;
6. Ліси, чагарники і луки;
7. Тундра.

Відповідно до карти тайлів та карт кожного континенту було обрано 8 тайлів для оцінки точності. Загалом, 9 типів земного покриття та 150000 тестових прикладів було оцінено з точки зору точності. Загальна точність GlobeLand30-2010 досягає 83.51%. Індикатор Каппа дорівнює 0,78.

GlobeLand30-2010 набір даних має такі характеристики системи координат:

- система координат: WGS84;
- референц-еліпсоїд: WGS84;
- проекція: UTM;
- зонування: 6°.

GlobeLand30-2010 продукт організований в тайлах:

- 60° Пн. – 60 Пд.: 5° (широта) x 6° (довгота);
- 60-80° Пн. і 60-80° Пд: 5° (широта) x 12° (довгота).

Кількість тайлів карти GlobeLand30-2010 є 853.

Тайлові дані GlobeLand30-2010 складаються з п'яти частин, а саме файлу результату класифікації, файлу інформації про координати, індексного картографічного файлу класифікації зображень, файлу метаданих та ілюстративного файлу. Серед них:

- файл класифікації результату: відповідає файлу зберігання інформації про класифікацію земного покриття;

– файл інформації про координати відповідає файлу запису координатної інформації результатів класифікації даних;

– індексний картографічний файл класифікації відповідає файлу векторного шару запису діапазону і часу кожної сцени класифікаційних зображень;

– файл метаданих відповідає файлу запису інформації про метадані результатів класифікації;

– ілюстративний файл відповідає файлу, що описує результат класифікації даних і пояснення авторських прав на дані.

Система класифікації включає в себе десять видів земного покриву, а саме сільськогосподарські землі, ліси, луки, чагарник, водно-болотні угіддя, водойми, тундру, урбанізовані території, пустелі, льодовий покрив. Ілюстрація кожного типу наведена в табл. 1.

Підготовка даних для території України складалась з таких етапів:

– завантаження тайлів, що покривають територію України;

– обрізання країв тайлів за допомогою векторних масок;

– злиття мозаїки;

– обрізання мозаїки по контуру державного кордону України;

– експорт даних у формат GeoTIFF.

На наступному етапі необхідно сформувати геостатистичні дані для адміністративних одиниць України на різних територіальних рівнях. Перевагою такого підходу є можливість формування статистичних показників для різних територіальних одиниць або їх поєднань: річкові басейни, природні зони та ін.

Формування геостатистичних даних відбувається з використанням операцій зональної статистики, які реалізовано в модулі Spatial Analyst ПЗ ArcGis. Для обробки даних якісного розподілу класифікаційних ознак підходить інструмент «Зональна статистика в таблицю», який дозволяє резюмувати значення растру у межах зон іншого набору даних.

Таблиця 1 – Класифікація GlobeLand30

Код	Тип	Зміст
10	Сільськогосподарські землі	Землі, які використовуються для сільського господарства, багаторічних насаджень, в тому числі рисових полів, зрошуваних сільгоспугідь, овочівництва, фруктових садів
20	Ліс	Землі, які вкриті лісовою рослинністю з покривом більше 30%, в тому числі листяні і хвойні ліси, і рідколісся з покриттям 10 – 30% і т.д.
30	Луки	Землі, які покриті натуральною трав'янистою рослинністю з покриттям більше 10%, і т.д.
40	Деревно-чагарникова рослинність	Землі, які вкриті чагарниками з покриттям більше ніж 30%, в тому числі листяні і вічнозелені чагарники і пустельний степ з покриттям більше 10%, і т.д.
50	Водно-болотні угіддя	Землі, які вкриті водно-болотною рослинністю і водоймами, в тому числі внутрішні болота, заболочені озера, заплави річок з водно-болотними угіддями, ліси / чагарники водно-болотних угідь, торф'яні болота і т.д.
60	Водні об'єкти	Водні об'єкти на суші, в тому числі річки, озера, водосховища, ставки для розведення риби і т.д.
70	Тундра *	Землі вкриті лишайником, мохом, витривалою багаторічною трав'янистою рослинністю і чагарниками в полярних регіонах, в тому числі чагарникової тундри, трав'янистих тундрах і т.д.
80	Урбанізовані території	Землі модифіковані діяльністю людини, в тому числі всі види житла, промислові і гірничодобувні галузі, транспортні комунікації, внутрішні міські зелені зони і водойми і т.д.
90	Пустельні території *	Земля з рослинним покривом менше 10%, в тому числі пустелі, піщані поля, Гобі, голі скелі і т.д.
100	Постійний сніговий або льодовий покрив*	Земля вкрита вічним снігом, льодовики і крижаний покрив

* – відсутність даного типу земного покриву в Україні або наявність у кількості менше 1 %

Зональні інструменти дозволяють виконувати аналіз, вихідні дані якого є результатом обчислень, виконаних на всіх комірках, що належать кожній вхідній зоні. Зона може бути визначена як область однакових значень, але вона також може складатися з декількох відокремлених елементів або регіонів, всі з яких мають одне значення. Зони можуть задаватися растром або наборами класів векторних об'єктів. Растр повинен мати тип ціле число, а просторові об'єкти повинні мати цілочисельне або строкове поле атрибутів.

Деякі зональні інструменти оцінюють кількість певних властивостей геометрії, або форми, вхідних даних зони, і їм не потрібні інші вхідні дані. Інші зональні інструменти використовують вхідні дані зони для визначення місць розташування, для яких будуть обчислюватися інші параметри, наприклад, статистика, площі або частота значень. Існує також зональний інструмент, який використовується для заповнення зазначених зон мінімальними значенням, що знаходяться на межі зони.

Отримані значення площ можна використовувати для оцінки розподілу фактичних площ землекористування, а також для дослідження співвідношення між різними типами землекористування у відсотках, що потребує нормалізації даних відносно загальної площі адміністративної одиниці.

Математичною основою моделі є штучна нейронна мережа на основі карт самоорганізації Кохонена (рис. 3).

Процедуру роботи карт самоорганізації описано в роботах Т. Кохонена, С. Хайкіна, Ф. Уоссермена та інших.

Опишемо процедуру побудови даної нейронної мережі. Вхідний простір даних $X \subset \mathcal{R}^n$ проєктується найчастіше на двовимірну решітку вузлів нейронів. Кожному вузлу i , $i = \overline{1, k}$, ставиться у відповідність параметричний вектор моделі

$$m_i = \{\mu_{1i}, \mu_{2i}, \dots, \mu_{ni}\} \in \mathcal{R}^n,$$

де k – кількість вузлів нейронів. У той же час вхідний вектор $x = \{x_1, x_2, \dots, x_n\} \in X \subset \mathcal{R}^n$ пов'язаний одночасно з усіма нейронами через скалярні ваги μ_{ji} .

Перед початком проведення навчання нейронної мережі всі ваги μ_{ji} ініціалізуються за допомогою генератора випадкових величин, але слід зазначити, що існують й інші методи.

Після ініціалізації мережі запускається так званий процес конкуренції, який полягає у визначенні найбільш подібних нейронів до векторів вхідних даних. Найбільш поширеним підходом до даного процесу є обчислення Евклідової відстані:

$$\|x - m_i\| = \sqrt{\sum_{j=1}^n (x_j - \mu_{ji})^2}, \quad i = \overline{1, k}, \quad j = \overline{1, n}. \quad (1)$$

Після подання вхідних векторів на входи нейронів, відбувається змагання нейронів шару Кохонена за правилом «переможець отримує все», згідно якого обирається нейрон-переможець, що є найближчим до вхідного вектора:

$$c = \arg \min_{i=1, k} \{\|x - m_i\|\}. \quad (2)$$

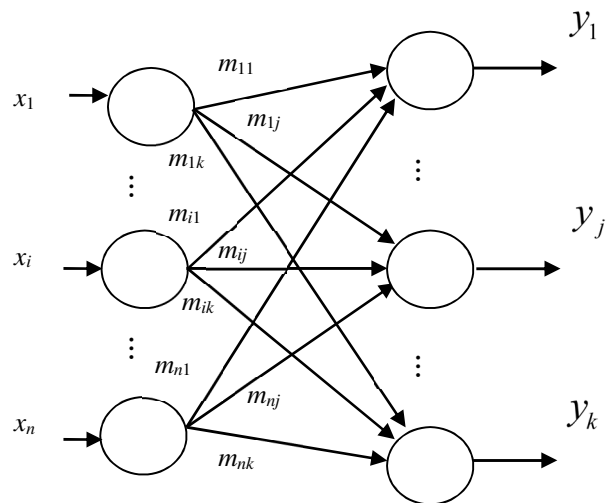


Рисунок 3 – Шар нейронів Кохонена

Таблиця 2 – Інструменти для роботи з зонами поділені на дві категорії з метою застосування

Категорія інструментів для роботи з зонами	Інструменти
Інструменти, що працюють з формами зони	Зональна геометрія (Zonal Geometry) Зональна геометрія в таблицю (Zonal Geometry As Table)
Інструменти, що працюють з атрибутами зони	Зональна статистика (Zonal Statistics) Зональна статистика в таблицю (Zonal Statistics as Table)
Інструменти, що визначають частку класів в межах зони	Таблиця площ (Tabulate Area)
Інструменти, що визначають розподіл частот значень одного вхідного растру в зонах, визначених іншим растром	Зональна гістограма (Zonal Histogram)
Інструменти, що заповнюють зазначені зони	Заповнення зон (Zonal Fill)

Після визначення нейрона-переможця здійснюється визначення околу його впливу на сусідні нейрони. Нейрон-переможець знаходиться в центрі топологічного околу. У разі збудження збудженні, він впливає на просторово близькі до нього нейрони, проте даний вплив зменшується зі збільшенням відстані між нейронами. Даний процес приводить до визначення топологічного сусідства нейронів.

На заключному етапі побудови карти самоорганізації відбувається процес синаптичної адаптації. Даний процес полягає у процедурі корегування вектора синаптичних ваг нейрона-переможця та його сусідів відповідно до вектора вхідного показника:

$$m_i(t+1) = m_i(t) + v(t)h_{ci}(t)[x(t) - m_i(t)], \quad (3)$$

де $v(t)$ – параметр швидкості навчання;

$h_{ci}(t)$ – функція топологічного сусідства з центром у нейроні-переможці c .

Результат налаштування такого типу нейронних мереж можна представити у вигляді карти Кохонена (рис. 4).

Опишемо властивості такої нейронної мережі. Даний підхід є унікальним – поєднує в собі процедури проєктування, кластеризації та виявлення взаємозв'язків між даними. Головною особливістю такого типу нейронних мереж є навчання «без вчителя». Це означає, що для коректного налаштування синаптичних ваг не потрібний бажаний відгук мережі, а отже процес моделювання протікає без втручання суб'єктивної думки розробника.

Оскільки карти самоорганізації здійснюють стиснення інформації із збереженням найбільш важливих топологічних та/або метричних зв'язків

між вхідними елементами даних, можна також вважати, що з їх допомогою створюються абстракції (узагальнення) певного виду. Ці дві властивості карт самоорганізації, візуалізацію та узагальнення, можна використовувати у вирішенні таких складних задач, як аналіз, управління та передача інформації.

В якості вхідних даних використовували розподіл земель між класами землекористування для кожної статистичної одиниці. За такі одиниці були обрані адміністративні райони України. Таким чином на вхід було подано 8 параметрів, які характеризують земний покрив України за виключенням пустельних районів та постійного снігового покриву. Відповідно до кількості вхідних параметрів було обрано кількість вихідних кластерів, припускаючи, що кожний додатковий клас вносить додаткове збурення та може виступати центром нового кластеру. В результаті було отримано 8 розподілів для кожного параметра та загальну кластеризацію на основі карт самоорганізації Кохонена (рис. 4).

Проведення кластеризації дозволило розрізнити кластери у географічному просторі шляхом розфарбування адміністративних районів за кольорами якісного фону на основі належності до відповідного кластеру. Як видно з рис. 3 значна кількість районів має сусідів, які належать до того ж кластеру, що показує зональність у розподілі кластерних ознак. Це, звичайно, закономірно, оскільки в основі аналізу покладено співвідношення між різними типами земного покриву, що в свою чергу залежить від таких зональних факторів у розподілі, як кліматичні ознаки та рослинний покрив. А зональними факторами в цьому випадку виступають антропогенно змінені території, які додають мозаїчність у просторовий розподіл кластерів (рис. 5).

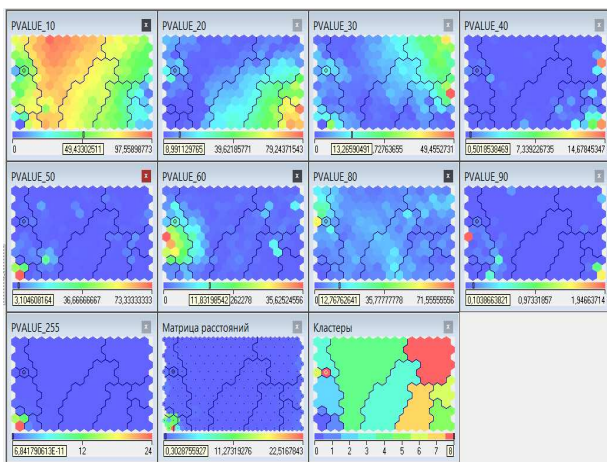


Рисунок 4 – Карти самоорганізації Кохонена на основі даних про земний покрив

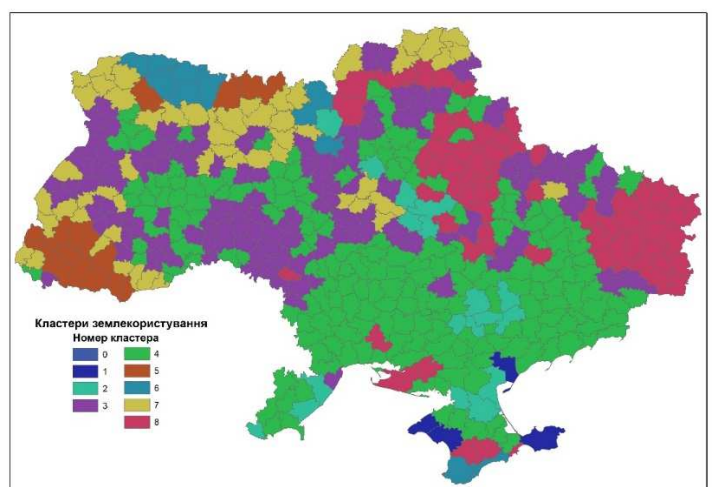


Рисунок 5 – Кластеризація адміністративних районів на основі даних про земний покрив

Кластер 0 характеризується вологими природними ландшафтами з переважанням лісових та болотних угідь.

Кластер 1 характеризується пустельними степовими ландшафтами півдня країни зі значним відсотком антропогенного втручання.

Кластер 2 визначається наявністю значних водних об'єктів у межах адміністративних одиниць, які у більшості випадків належать до водосховищ Дніпровського каскаду. Таким чином, визначним видом функціонування даних районів є водне господарство.

Кластер 3 характерний для лісостепової зони і характеризується рівномірним співвідношенням між такими групами земного покриву, як ліси, сільськогосподарські землі, урбанізовані території, лучна рослинність. Спеціалізація кластерів: сільськогосподарське виробництво та переробка, машинобудування, рекреація та охорона природи.

Кластер 4 зосереджений здебільшого на півдні України та на Поділлі. Кластер характеризується домінуванням у структурі землекористування сільськогосподарських земель, зокрема рілля та сіножатей, а також сільських населених пунктів. Кластер має компакту структуру, яка залежить від сприятливих агрокліматичних умов.

Кластер 5 належить до зон поширення хвойних лісів та чагарників. Зосередження кластеру спостерігається у Карпатських горах та у Поліссі. Основні види діяльності пов'язані з лісовим господарством та екстенсивними формами тваринництва. Також активно розвивається рекреаційна діяльність та охорона природи.

Кластер 6 належить до кластерів що характеризуються високим різноманіттям форм землекористування з переважанням урбанізованих територій та лісових масивів.

Кластер 7 має у своєму складі переважання мішаних та листяних лісів, чагарників, луків з невеликим відсотком антропогенно змінених територій. Основні види діяльності пов'язані з лісовим господарством, тваринництвом, рекреацією.

Кластер 8 характеризується переважанням сільськогосподарських земель та урбанізованих територій з високим відсотком змінених ландшафтів.

Основними функціями цих кластерів є промислове виробництво та переробка сільськогосподарської продукції.

Висновки

Можна зазначити, що наявність глобального продукту аналізу земного покриву Globeland 30 надає можливість для здійснення ефективної політики в галузі регіонального управління та територіального планування з урахуванням реально існуючого розподілу типів землекористування. Продукт Globeland 30 створено на основі автоматизованої класифікації знімків Landsat TM і ETM+ та знімків з китайського супутника HJ-1. Роздільна здатність та геометрична точність Globeland 30 дорівнюють характеристикам знімків Landsat і становлять відповідно 30 та 75 м. За методом класифікації Globeland 30 дуже схоже з верхніми рівнями класифікації CORINE, що робить її використання інтегрованим для участі в європейських проєктах.

В ході дослідження було розроблено схему та методику обробки даних для задач дослідження території України, яка складається з етапів передпроцесінга даних та отримання зональної статистики. На основі отриманих статистичних даних побудовано кластери розподілу існуючих адміністративних районів України за співвідношенням основних типів земних покривів на основі використання одного з видів нейронних мереж – карт самоорганізації Кахонена.

У результаті аналізу отриманих кластерів встановлено можливість прикладного використання кластеризації за адміністративними одиницями для розробки напрямів, сценаріїв та планів розвитку території, з'ясування її сучасного стану використання та корегування цих показників у майбутньому.

Подальші дослідження пов'язані з аналізом даних землекористування на основі інших показників різноманіття та співвідношення структур, кластеризації адміністративних одиниць з використанням методів машинного навчання.

Список літератури

1. Бродский Л. Проект INTAS по разработке автоматизированной технологии классификации земных покрытий: научные задачи, основные результаты и перспективы [Текст] / Л. Бродский, Е. И. Бушуев, В. И. Волошин, А. А. Козлова, О. И. Паршина, М. А. Попов, В. И. Саблина, А. И. Сахацкий, А. В. Сиротенко, Т. Сокуп, С. А. Станкевич, А. Г. Тарарико // *Космична наука і технологія*. 2009. – Т. 15. – № 2. – С. 36–48.
2. Попов М. А. Использование полного набора нормализованных межканальных индексов многоспектральных космических изображений при классификации покрытий ландшафта [Текст] / М. А. Попов, С. А. Станкевич, А. И. Сахацкий, А. А. Козлова // *Уч. зап. Таврического нац. ун-та им. В. И. Вернадского*. – 2007. – 20 (59), № 1. – С. 175–182.
3. Bossard M., Feranec J., Otahel J. The revised and supplemented Corine Land Cover nomenclature // *Techn. Rept EEA*. – 2000. – N 38. – 110 p.

4. Brian O'Connor, Cristina Secades, Johannes Penner, Ruth Sonnenschein, Andrew Skidmore, Neil D. Burgess & Jon M. Hutton *Earth observation as a tool for tracking progress towards the Aichi Biodiversity Targets // Remote Sensing in Ecology and Conservation* – у John Wiley & Sons Ltd, 2015. – P. 19 – 27
5. Buttner G., Feranec J., Jaffrain G., et al. *The CORINE Land Cover 2000 Project // EARSeL Proceedings* 3(3). – 2004. – P. 331 – 346.
6. Defourny et al., P. Defourny, L. Schouten, S. Bartalev, S. Bontemps, P. Caccetta, A.J.W. De Wit, C. Di Bella, et al. *Accuracy Assessment of a 300 M Global Land Cover Map: the GlobCover Experience* (2008).
7. Fritz et al., S. Fritz, I. McCallum, C. Schill, C. Perger, R. Grillmayer, F. Achard, F. Kraxner, M. Obersteiner *Geo-Wiki.Org: the use of crowdsourcing to improve global land cover Remote Sens.*, 1 (3) (2009), pp. 345–354 <http://dx.doi.org/10.3390/rs1030345>
8. Jan Feraneca, Gerard Hazeub, Susan Christensenc, Gabriel Jaffraind, *Corine land cover change detection in Europe (case studies of the Netherlands and Slovakia) // Land Use Policy*, Volume 24, Issue 1, January 2007, Pages 234 – 247.
9. National Geomatics Center of China, *2014 National Geomatics Center of China 30 M Global Land Cover Data Product (GlobeLand30)* (2014).
10. Pashynska Nataliia *A decision tree in a classification of fire hazard factors / Nataliia Pashynska, Vitaliy Snytyuk, Viktor Putrenko, Andriy Musienko // Eastern-European journal of enterprise technologies.* – Vol 5. – NO 10 (83). – 2016. – pp. 32– 37.
11. *The thematic accuracy of Corine land cover 2000. Assessment using LUCAS (land use/cover area frame statistical survey) // Techn. Rept. EEA.* – 2006. – N 7. – 85 p. – (www.eea.europa.eu).

Стаття надійшла до редакції 31.07.2017

Рецензент: д-р техн. наук, проф. А.А. Лященко, Київський національний університет будівництва та архітектури, Київ

Путренко Виктор Валентинович

Кандидат географических наук, докторант, orcid.org/0000-0002-0239-9241

Национальный технический университет Украины «Киевский политехнический институт им. Игоря Сикорского», Киев

Пашинская Наталия Николаевна

Кандидат географических наук, старший научный сотрудник, orcid.org/0000-0002-0133-688X

Киевский национальный университет им. Тараса Шевченко, Киев

ИНТЕЛЛЕКТУАЛЬНЫЙ АНАЛИЗ ЗЕМНОГО ПОКРОВА ТЕРРИТОРИИ УКРАИНЫ НА ОСНОВЕ ДАННЫХ GLOBELAND 30

Аннотация. Рассмотрен анализ и типизация земного покрова территории Украины для целей государственного территориального управления и администрирования с использованием данных продукта глобальной классификации земного покрова Globeland 30. Для проведения типизации использовался набор методов ГИС-анализа, в частности обработки растровых наборов данных, получения тайловых мозаик и конвертации, обработки векторных данных путем переклассификации и зональной статистики, методы кластеризации, в частности метод построения карт самоорганизации Кохонена для целей кластеризации и типизации административных единиц. На основе этих данных проведена типизация административных единиц и получено распределение на 9 кластеров, которые отражают основные типы землепользования в Украине. Таким образом, получены группы районов, которые по своему территориальному положению, специализации и особенностям развития имеют подобные структуры землепользования. Полученные результаты типизации являются основой для формирования и корректировки программ территориального развития на уровне административных единиц, контроля за природоохранной, сельскохозяйственной и градостроительной деятельностью.

Ключевые слова: интеллектуальный анализ; спутниковые снимки; земной покров; землепользование; ГИС-анализ; карты самоорганизации Кохонена; кластеризация

Putrenko Viktor

PhD (Geograf.), Doctoral student, orcid.org/0000-0002-0239-9241

National technical university of Ukraine "Igor Sikorsky Kyiv polytechnic institute

Pashynska Nataliia

PhD (Geograf.), Senior Researcher, orcid.org/0000-0002-0133-688X

Taras Shevchenko National University of Kyiv

DATA MINING OF LAND COVER IN UKRAINE BASED ON GLOBELAND 30 DATA

Abstract. Analysis and typology of land cover of the territory of Ukraine for the goal of state territorial management and administration using the land cover classification product data global Globeland 30 was considered in work. For the typing was

used a set of GIS analysis methods, in particular the processing of raster data sets, obtain tile mosaics and conversion processing of vector data through reclassification and zonal statistics, methods of clustering, in particular the method of constructing Kohonen self-organizing map for the purpose of clustering and typing of administrative units. On the basis of these data typing performed administrative units and received distribution on 9 clusters that represent the main types of land use in Ukraine. Thus, a group of areas of their territorial location, specialization and features of land have similar structures were received. The results of typing are the basis for development and adjustment programs of territorial development at administrative units, control of environmental, agricultural, urban activities.

Keywords: Data mining; satellite imagery; land cover; land use; GIS analysis; Kohonen self-organizing map; clustering

References

1. Brodskyy, L., Bushuev, E.Y., Voloshyn, V.Y., Kozlova, A.A., Parshyna, O.Y., Popov, M.A., Sablyna, V.Y., Sakhatsky, A.Y., Syrotenko, A.V., Soukup, T., Stankevych, S.A., Tararyko, A.H. (2009). INTAS project on the development of an automated technology for the classification of terrestrial coverages: scientific tasks, main results and prospects. *Space science and technology*, 15, 2, 36–48.
2. Popov, M.A., Stankevych, S.A., Sakhatsky, A.Y., Kozlova, A.A. (2007). The use of a complete set of normalized interchannel indices of multispectral cosmic images in the classification of landscape coverings. *Scientific notes of the Vernadsky Tavrda National University*, 20 (59), 1, 175–182.
3. Bossard, M., Feranec, J., Otahel, J. (2000). *The revised and supplemented Corine Land Cover nomenclature*. Techn. Rept EEA, 38, 110.
4. O'Connor, Brian, Secades, Cristina, Penner, Johannes, Sonnenschein, Ruth, Skidmore, Andrew, Burgess, Neil D. & Hutton, Jon M. (2015). *Earth observation as a tool for tracking progress towards the Aichi Biodiversity Targets*. *Remote Sensing in Ecology and Conservation*. John Wiley & Sons Ltd, 19–27.
5. Buttner, G., Feranec, J., Jaffrain, G., et al. (2004). *The CORINE Land Cover 2000 Project*. *EARSel Proceedings*, 3(3), 331–346.
6. Defourny et al. (2008). *Accuracy Assessment of a 300 M Global Land Cover Map: the GlobCover Experience*. P. Defourny, L. Schouten, S. Bartalev, S. Bontemps, P. Caccetta, A.J.W. De Wit, C. Di Bella, et al.
7. Fritz et al. (2009). S. Fritz, I. McCallum, C. Schill, C. Perger, R. Grillmayer, F. Achard, F. Kraxner, M. Obersteiner *Geo-Wiki.Org: the use of crowdsourcing to improve global land cover Remote Sens.*, 1 (3), 345–354 <http://dx.doi.org/10.3390/rs1030345>
8. Feraneca, Jan, Hazeub, Gerard, Christensenc, Susan, Jaffraind, Gabriel. (2007). *Corine land cover change detection in Europe (case studies of the Netherlands and Slovakia)*. *Land Use Policy*, 24, Issue 1, 234–247.
9. *National Geomatics Center of China, 2014 National Geomatics Center of China 30 M Global Land Cover Data Product (GlobeLand30) (2014)*
10. Pashynska, Nataliia, Snytyuk, Vitaliy, Putrenko, Viktor, Musienko, Andriy. (2016). *A decision tree in a classification of fire hazard factors*. *Eastern-European journal of enterprise technologies*, 5, 10 (83), 32–37.
11. *The thematic accuracy of Corine land cover 2000. Assessment using LUCAS (land use/cover area frame statistical survey) // Techn. Rept. EEA, 2006, 7, 85. – (www.eea.europa.eu)*.

Посилання на публікацію

- APA Putrenko, Viktor & Pashynska, Nataliia. (2017). *Data mining of land cover in Ukraine based on Globeland 30 data. Management of Development of Complex Systems*, 31, 117 – 125.
- ГОСТ Пупренко В.В. Інтелектуальний аналіз земного покриття території України на основі даних Globeland 30 [Текст] / В.В. Пупренко, А.В. Третьяк // *Управління розвитком складних систем*. – 2017. – № 31. – С. 117 – 125.