
ЗЕМЛЕРОБСТВО, РОСЛИННИЦТВО, ОВОЧІВНИЦТВО ТА БАШТАННИЦТВО

AGRICULTURE, CROP PRODUCTION,
VEGETABLE AND MELON GROWING

УДК 528.88: 633.1: 633.34: 633.854.78: 631.67
DOI <https://doi.org/10.32782/2226-0099.2024.136.2.1>

КЛАСИФІКАЦІЯ ЗРОШУВАНИХ І НЕЗРОШУВАНИХ ПОСІВІВ ПШЕНИЦІ ОЗИМОЇ, КУКУРУДЗИ, СОЇ ТА СОНЯШНИКУ НА ОСНОВІ ДАНИХ АЕРОКОСМІЧНОГО МОНІТОРИНГУ

Лиховид П.В. – д.с.-г.н.,

старший науковий співробітник, відділ зрошувального землеробства
та декарбонізації агроєкосистем,

Інститут кліматично орієнтованого сільського господарства
Національної академії аграрних наук України

Грановська Л.М. – д.е.н., професор,

член-кореспондент Національної академії аграрних наук України,
завідувач відділу зрошувального землеробства та декарбонізації агроєкосистем,

Інститут кліматично орієнтованого сільського господарства
Національної академії аграрних наук України

Дані дистанційного зондування Землі є перспективними для розпізнавання зрошуваних і незрошуваних земель, що є важливим і актуальним завданням сучасної аграрної науки, оскільки дозволить виконувати оперативну динамічну оцінку реальних масштабів впровадження зрошення та забезпечити раціональне використання водних ресурсів в умовах наростаючого їх дефіциту. Для класифікації зрошуваних і незрошуваних посівів пшениці озимої, кукурудзи на зерно, соняшнику та сої, як основних культур Півдня України, було залучено дані щодо часової серії нормалізованого диференційного вегетаційного індексу, розрахованого за комбінованими аерокосмічними знімками satelіtіv Landsat-8 і Sentinel-2 з роздільною здатністю 250 м. Узагальнені дані по 200 рендомізовано обраних та закріплених полях (1200 вхідних незалежних змінних) було проаналізовано методами мульти-класового лінійного та канонічного дискримінантного аналізу, та побудовано відповідні класифікуючі функції за рівня достовірності 95%. Математичні розрахунки та статистичний аналіз дослідних даних виконували у статистичному пакеті BioStat. Загальний коефіцієнт кореляції склав 0,81 за коректності класифікації 91,5%. Варто відзначити істотне коливання точності класифікації за досліджуваними культурами. Максимальну точність класифікації досягнуто для сої (відсоток коректності – 98,0% за коефіцієнта кореляції 0,94), а мінімальну – для пшениці озимої (відсоток коректності – 86,0% за коефіцієнта кореляції 0,80). В цілому результати математичної оцінки та статистичного аналізу засвідчили високу точність і якість класифікації зрошуваних і незрошуваних земель за даними розрахункового нормалізованого диференційного вегетаційного індексу

з аерокосмічних знімків. У подальшому розроблені дискримінантні функції буде залучено в систему автоматизованої ідентифікації посівів.

Ключові слова: дискримінантний аналіз, зрошення, математичний аналіз, нормалізований диференційний вегетаційний індекс, супутникові знімки.

Lykhovyd P.V., Hranovska L.M. Classification of the irrigated and non-irrigated crops of winter wheat, corn, soybeans and sunflower by the means of aerospace monitoring

Remote sensing data are promising for the recognition of irrigated and non-irrigated lands, which is an important and urgent task of modern agricultural science, as it will allow to perform an operational dynamic assessment of the real scale of irrigation implementation and ensure the rational use of water resources in the conditions of increasing water scarcity. For the classification of irrigated and non-irrigated crops of winter wheat, grain corn, sunflower and soybeans, as the main crops of Southern Ukraine, data on the time series of the normalized difference vegetation index, calculated using the combined aerospace images retrieved from Landsat-8 and Sentinel-2 satellites with a resolution of 250 m, were applied. Generalized data on 200 randomly selected and fixed fields (1200 input independent variables) were analyzed using the methods of multiclass linear and canonical discriminant analysis, and corresponding classification functions were built up at the 95% confidence level. Mathematical calculations and statistical analysis of the experimental data were performed in the statistical package BioStat. The generalized correlation coefficient was 0.81 with a classification accuracy of 91.5%. It is worth noting the significant fluctuation of the accuracy of the classification depending on the studied crops. The maximum classification accuracy was achieved for soybeans (percentage of correctness – 98.0% with a correlation coefficient of 0.94), and the minimum – for winter wheat (percentage of correctness – 86.0% with a correlation coefficient of 0.80). In general, the results of the mathematical evaluation and statistical analysis proved the high accuracy and quality of the classification of irrigated and non-irrigated lands based on the data of the calculated normalized difference vegetation index from aerial photographs. In the future, the discriminant functions developed will be included in the automated crop identification system.

Key words: discriminant analysis, irrigation, mathematical analysis, normalized difference vegetation index, aerospace imagery.

Постановка проблеми. Супутниковий моніторинг є перспективною технологією для ідентифікації, динамічного моніторингу, здійснення оперативного управління та контролю за станом сільськогосподарських земель. Застосування даних аерокосмічного моніторингу відкриває нові можливості для дистанційної ідентифікації зрошуваних земель, що є важливим для оцінки реального масштабу впровадження зрошувальних меліорацій, контролю за станом зрошуваних сільськогосподарських угідь, їх географічної локалізації та прив'язки з метою наступного картування, а також оптимізації використання водних ресурсів [1]. Враховуючи високу актуальність даного напрямку досліджень та практичну відсутність первинних вітчизняних наукових напрацювань із цієї проблеми, нами було вирішено за доцільне поставити на вивчення та розробити методологію застосування даних дистанційного зондування Землі для розпізнавання зрошуваних і незрошуваних посівів основних сільськогосподарських культур, вирощуваних на Півдні України, а саме пшениці озимої, кукурудзи на зерно, сої та соняшнику, побудовану на основі математичного дискримінантного аналізу.

Аналіз останніх досліджень. Проблематиці розпізнавання зрошуваних земель за даними супутникового моніторингу в останні десятиліття присвячено достатню увагу, втім, остаточного універсального вирішення цього питання донині не знайдено, а всі методичні підходи апробуються та використовуються локально в тих умовах і агроекологічних зонах, для яких вони були розроблені. На сьогодні наукове вирішення проблеми розпізнавання зрошуваних і незрошуваних земель знаходиться на проміжному етапі, і потребує подальшої наукової розробки та уніфікації. Загальноновизнаним світовою науковою спільнотою є лише той факт, що розрахункові вегетаційні індекси є ефективним вхідним параметром

для виконання класифікації сільськогосподарських земель. Крім того, коректність розпізнавання значною мірою залежатиме від видів сільськогосподарських культур, вирощуваних на зрошенні [2]. Нижче наведено кілька прикладів останніх наукових досліджень, пов'язаних із тематикою дистанційного розпізнавання та наступного картування зрошуваних і незрошуваних земель.

Так, дослідження, виконане для умов Гани засвідчило про можливість класифікації земель сільськогосподарського призначення із використанням даних щодо величини нормалізованого диференційного вегетаційного індексу (NDVI). Науковці використали алгоритм дерева ухвалення рішень для нечіткого набору вхідних даних. У результаті, коректність класифікації зрошуваних земель коливалася в межах 67–93% [3]. Подібний алгоритм було використано і у більш масштабному дослідженні глобального характеру з вивчення структури та розподілу площ зрошуваних земель у світі за період 1999–2012 рр. на основі часової серії даних щодо величини NDVI; точність розпізнавання коливалася залежно від джерела даних у межах 82–88% для незрошуваних, і 83–92% для зрошуваних земель, відповідно [4].

Дослідження, виконані в Англії щодо можливості класифікації зрошуваної та незрошуваної картоплі на основі даних дистанційного зондування Землі, виявили, що в умовах вологого клімату закономірності та сезонна динаміка вегетаційних індексів на зрошуваних і незрошуваних землях є майже ідентичною, а тому ефективно вирішення проблеми такої класифікації досі не винайдено. Таким чином, залучення аерокосмічних зйомок є раціональним лише в умовах посушливих, де простежується чіткий контраст між продуктивністю та відповідними спектральними характеристиками посівів сільськогосподарських культур [5]. Саме такі умови в останні десятиліття сформувалися на території України внаслідок кліматичних трансформацій на тлі глобального потепління [6], тому перспективність залучення даних супутникового моніторингу посівів для класифікації зрошуваних і незрошуваних земель не викликає сумніву.

Варто відзначити, що більшість класифікацій здійснювалася із залученням алгоритмів дерева ухвалення рішень, а у якості вхідного параметра найбільш часто застосовували часову серію NDVI [7]. Також успішно було застосовано для подібної задачі ансамблевий метод машинного навчання (random forest) [8], а традиційні методи дискримінантного аналізу, логістичної регресії, а також штучні нейронні мережі представлені значно менше. Враховуючи високі вимоги до обчислювальних потужностей, а також відносну закритість алгоритмів штучних нейронних мереж, що знижує їх універсальність не дивлячись на ряд переваг у точності [9], нами було обрано традиційний дискримінантний аналіз, який у якості класифікуючої надає лінійну канонічну функцію. Логістичний аналіз, як вважається, має вищу точність, але суттєва різниця, як уже було доведено працями окремих науковців, досягається лише за умови нечітких класів [10–12].

Постановка завдання. Завданням представленого дослідження було проаналізувати можливості залучення даних дистанційного зондування Землі, а саме величини NDVI, для класифікації та розпізнавання зрошуваних і незрошуваних земель в умовах Півдня України для таких сільськогосподарських культур як пшениця озима, кукурудза на зерно, сояшник і соя, та розробити методологію класифікаційного аналізу на основі результатів первинних досліджень.

Виклад основного матеріалу дослідження. Аналіз можливості застосування часової серії NDVI для розпізнавання зрошуваних і незрошуваних посівів основних сільськогосподарських культур, вирощуваних на Півдні України, виконували за даними з комбінованих супутникових знімків сателітів Landsat-8

і Sentinel-2 з роздільною здатністю 250 м, виключаючи з часової серії спотворені, неякісні знімки та знімки з високим відсотком хмарності. Часову серію вегетаційного індекса для рендомізовано обраних і потім закріплених полів кожної з досліджуваних культур (пшениця озима, кукурудза на зерно, соя, соняшник) узагальнювали за помісячними періодами (у часовий проміжок «травень – жовтень»). Всього було проаналізовано по 50 полів кожної культури (25 зрошуваних і 25 незрошуваних), розташованих переважно у Херсонській та Миколаївській областях, за період 2018 року. Генералізована вибірка даних склала 200 полів, загальна кількість вхідних даних щодо величини NDVI – 1200. Математичний аналіз виконували за алгоритмами мультикласового лінійного дискримінантного аналізу (MLDA) та канонічного дискримінантного аналізу (CDA) [13]. Розрахунки виконували у статистичному пакеті BioStat v.7 за $P < 0,05$. За результатами математичної обробки даних було побудовано канонічну дискримінантну функцію для розпізнавання зрошуваних і незрошуваних земель в цілому, а також для кожної окремої культури.

У результаті виконання математичних розрахунків було пораховано канонічний коефіцієнт кореляції, який склав 0,81, що свідчить про високу статистичну якість і адекватність дискримінантної функції. Розрахунок таких статистичних показників як лямбда Уїлкса (0,34) і слід Пілая (0,66) дозволили з упевненістю відхилити нульову гіпотезу (про неможливість класифікації). Додатковим свідченням на користь можливості класифікації зрошуваних і незрошуваних земель за вхідним набором місячної часової серії NDVI є величина критерія $\chi^2_{\text{факт}} = 207,97$ за числа ступенів свободи 6 (контрольне значення критерію становить $\chi^2 = 12,60 < \chi^2_{\text{факт}}$), що дозволяє повністю спростувати нульову гіпотезу.

Об'єднана варіаційно-коваріаційна матриця, а також повна матриця класифікації наведені у табл. 1. Канонічні коефіцієнти та канонічна структура класифікаційної функції наведені у табл. 2.

Таблиця 1

Об'єднана варіаційно-коваріаційна матриця та повна матриця класифікації зрошуваних і незрошуваних земель за даними NDVI

Об'єднана варіаційно-коваріаційна матриця класифікації						
Змінна (NDVI)	травень	червень	липень	серпень	вересень	жовтень
травень	0,0248	0,0323	-0,0020	-0,0270	-0,0223	-0,0042
червень	0,0323	0,0474	-0,0021	-0,0399	-0,0317	-0,0059
липень	-0,0020	-0,0021	0,0103	0,0029	-0,0034	-0,0017
серпень	-0,0270	-0,0399	0,0029	0,0443	0,0307	0,0052
вересень	-0,0223	-0,0317	-0,0034	0,0307	0,0385	0,0097
жовтень	-0,0042	-0,0059	-0,0017	0,0052	0,0097	0,0058
Повна матриця класифікації						
Змінна (NDVI)	травень	червень	липень	серпень	вересень	жовтень
травень	0,0252	0,0323	-0,0015	-0,0246	-0,0199	-0,0039
червень	0,0323	0,0473	-0,0019	-0,0389	-0,0307	-0,0058
липень	-0,0015	-0,0019	0,0108	0,0055	-0,0008	-0,0013
серпень	-0,0246	-0,0389	0,0055	0,0555	0,0422	0,0069
вересень	-0,0199	-0,0307	-0,0008	0,0422	0,0501	0,0115
жовтень	-0,0039	-0,0058	-0,0013	0,0069	0,0115	0,0061

Таблиця 2

Канонічні коефіцієнти та канонічна структура функції класифікації зрошуваних і незрошуваних земель за даними NDVI

Змінна (NDVI)	Канонічні коефіцієнти		
	Звичайні	Стандартизовані	Повна структура
травень	5,6108	0,8843	0,1639
червень	4,3191	0,9406	0,0456
липень	3,7303	0,3784	0,2913
серпень	5,1812	1,0900	0,5610
вересень	5,7194	1,1223	0,5988
жовтень	-2,6224	-0,2004	0,2581

Функція класифікації для зрошуваних і незрошуваних земель півдня України наведена у табл. 3. Матриця класифікації (результуюча) представлена у табл. 4.

Таблиця 3

Функція класифікації зрошуваних і незрошуваних земель за даними NDVI

Група	V	VI	VII	VIII	IX	X	константа
зрошувані	28,1418	61,3133	52,9820	54,9718	34,2783	21,5056	-46,7715
незрошувані	12,7304	49,4499	42,7358	40,7402	18,5686	28,7086	-25,9281

Таблиця 4

Матриця класифікації зрошуваних і незрошуваних земель за даними NDVI

Група \ Передбачена	зрошувані	незрошувані	Всього	Відсоток коректності
зрошувані	91	9	100	91,0%
незрошувані	8	92	100	92,0%
Всього	99	101	200	91,5%
<i>N коректних = 183</i>				

Таким чином, встановлено, що за даними помісячної часової серії NDVI можна з доволі високою (91,5%) точністю виконати дистанційну класифікацію зрошуваних і незрошуваних посівів основних сільськогосподарських культур Півдня України.

Додатково надаємо функції класифікації (табл. 5), а також результати матриці класифікації зрошуваних і незрошуваних посівів для кожної досліджуваної культури окремо (табл. 6).

Встановлено, що максимальну (98,0%) точність класифікації канонічні функції забезпечують для посівів кукурудзи на зерно та сої, у той час як мінімальну (86,0%) – для пшениці озимої. При цьому відмічено певні міжгрупові відмінності. Так, наприклад, зрошувані та незрошувані посіви соняшнику ідентифікуються з однаковою точністю; для пшениці озимої та кукурудзи на зерно класифікуюча функція забезпечувала максимальну точність у розпізнаванні незрошуваних земель, у той час як для сої – навпаки, зрошуваних. Максимальний коефіцієнт кореляції для канонічної функції встановлено для сої, у той час як

мінімальний – для пшениці озимої. Ці особливості необхідно враховувати під час застосування розроблених дискримінантних функцій класифікації у наукових і практичних цілях.

Таблиця 5

Функції класифікації зрошуваних і незрошуваних посівів досліджуваних культур за даними NDVI

Група	V	VI	VII	VIII	IX	X	константа
пшениця озима							
зрошені	-9,7021	169,2668	-58,1099	69,8045	121,5592	219,9100	-71,8640
незрошені	-26,1876	154,4067	-66,0404	68,3065	114,6071	230,0972	-52,5441
соляшник							
зрошені	109,0168	23,6479	124,0220	111,3787	111,4712	29,3895	-94,4985
незрошені	91,7896	52,7838	100,7603	89,8808	88,7875	26,2191	-66,6667
кукурудза на зерно							
зрошені	33,7654	95,2327	78,4457	72,7962	36,1976	46,9554	-65,5002
незрошені	63,4961	78,7181	58,5888	50,8028	15,9501	39,4445	-36,8342
соя							
зрошені	69,8560	74,8651	109,4211	94,6726	53,6681	-25,5538	-79,6537
незрошені	82,9034	47,4025	81,1902	49,3747	25,5037	-4,4578	-34,8158

Таблиця 6

Матриці класифікації зрошуваних і незрошуваних земель за даними NDVI для пшениці озимої, соляшнику, кукурудзи на зерно та сої

Група \ Передбачена	зрошені	незрошені	Всього	Відсоток коректності
пшениця озима				
зрошені	20	5	25	80,0%
незрошені	2	23	25	92,0%
Всього	22	28	50	86,0%
<i>N коректних = 43; коефіцієнт кореляції 0,80</i>				
соляшник				
зрошені	24	1	25	96,0%
незрошені	1	24	25	96,0%
Всього	25	25	50	96,0%
<i>N коректних = 48; коефіцієнт кореляції 0,81</i>				
кукурудза на зерно				
зрошені	24	1	25	96,0%
незрошені	0	25	25	100,0%
Всього	24	26	50	98,0%
<i>N коректних = 49; коефіцієнт кореляції 0,88</i>				
соя				
зрошені	25	0	25	100,0%
незрошені	1	24	25	96,0%
Всього	26	24	50	98,0%
<i>N коректних = 49; коефіцієнт кореляції 0,94</i>				

Висновки та перспективи подальших досліджень. Результати досліджень засвідчили про можливість використання часової серії нормалізованого диференційного вегетаційного індексу та дискримінантних канонічних функцій для високоточної дистанційної класифікації зрошуваних і незрошуваних земель на Півдні України. Результати дослідження відкривають нові перспективи подальшого розширеного залучення даних аерокосмічної зйомки у вітчизняну аграрну науку і практику, а також будуть використані для розробки автоматизованої системи ідентифікації посівів.

СПИСОК ВИКОРИСТАНОЇ ЛІТЕРАТУРИ:

1. Pageot Y., Baup F., Inglada J., Baghdadi N., Demarez V. Detection of irrigated and rainfed crops in temperate areas using Sentinel-1 and Sentinel-2 time series. *Remote Sensing*. 2020. Vol. 12. No. 18. P. 3044.
2. Ozdogan M., Yang Y., Allez G., Cervantes C. Remote sensing of irrigated agriculture: Opportunities and challenges. *Remote Sensing*. 2010. Vol. 2. No. 9. P. 2274-2304.
3. Gumma M. K., Thenkabail P. S., Hideto F., Nelson A., Dheeravath V., Busia D., Rala A. Mapping irrigated areas of Ghana using fusion of 30 m and 250 m resolution remote-sensing data. *Remote Sensing*. 2011. Vol. 3. No. 4. P. 816-835.
4. Meier J., Zabel F., Mauser W. A global approach to estimate irrigated areas—a comparison between different data and statistics. *Hydrology and Earth System Sciences*. 2018. Vol. 22. No. 2. P. 1119-1133.
5. Shamal S. A. M., Weatherhead K. Assessing spectral similarities between rainfed and irrigated croplands in a humid environment for irrigated land mapping. *Outlook on Agriculture*. 2014. Vol. 43. No. 2. P. 109-114.
6. Lykhovyd P. Irrigation needs in Ukraine according to current aridity level. *Journal of Ecological Engineering*. 2021. Vol. 22. No. 8. P. 11-18.
7. Pervez M. S., Budde M., Rowland J. Mapping irrigated areas in Afghanistan over the past decade using MODIS NDVI. *Remote Sensing of Environment*. 2014. Vol. 149. P. 155-165.
8. Demarez V., Helen F., Marais-Sicre C., Baup F. In-season mapping of irrigated crops using Landsat 8 and Sentinel-1 time series. *Remote Sensing*. 2019. Vol. 11. No. 2. P. 118.
9. Chen L., Li S., Bai Q., Yang J., Jiang S., Miao Y. Review of image classification algorithms based on convolutional neural networks. *Remote Sensing*. 2021. Vol. 13. No. 22. P. 4712.
10. El-Habil A., El-Jazzar M. A comparative study between linear discriminant analysis and multinomial logistic regression. *An-Najah University Journal for Research-B (Humanities)*. 2013. Vol. 28. No. 6. P. 1525-1548.
11. Liong C. Y., Foo S. F. Comparison of linear discriminant analysis and logistic regression for data classification. *AIP Conference Proceedings*. 2013. Vol. 1522. No. 1. P. 1159-1165.
12. Dreiseitl S., Ohno-Machado L. Logistic regression and artificial neural network classification models: a methodological review. *Journal of Biomedical Informatics*. 2002. Vol. 35. No. 5-6. P. 352-359.
13. Klecka W. R. *Discriminant analysis*. Sage, 1980. 105 pp.