



Агроекологія, радіологія, меліорація

УДК 519.2:528.88:633.8

© 2025

МАТЕМАТИЧНЕ МОДЕЛЮВАННЯ ВРОЖАЙНОСТІ ОЛІЙНИХ КУЛЬТУР НА ПІВДНІ УКРАЇНИ ЗА ДАНИМИ АЕРОКОСМІЧНОГО МОНІТОРИНГУ

Р.А. Вожегова¹, П.В. Лиховид², Л.М. Грановська³

¹доктор сільськогосподарських наук, професор, академік НААН

²доктор сільськогосподарських наук

³доктор економічних наук, професор, член-кореспондент НААН

Інститут кліматично орієнтованого сільського господарства

Національної академії аграрних наук України

вул. Маяцька дорога, 24, смт Хлібодарське

Одеського р-ну Одеської обл., 67667, Україна

e-mail: ¹izz.biblio@ukr.net, ²pavel.likhovid@gmail.com, ³G_Ludmila15@ukr.net

ORCID: ¹0000-0002-3895-5633, ²0000-0002-0314-7644, ³0000-0001-7021-3093

Надійшла 03.02.2025

Мета. Розробити математичні моделі врожайності основних олійних культур, вирощуваних на Півдні України, з використанням даних аерокосмічного моніторингу їх посівів. Оцінити достовірність і точність моделей, визначити лімітувальні фактори та можливості їх науково-практичного використання. **Методи.** Дослідження виконано з використанням даних середньомісячної величини нормалізованого диференційного вегетаційного індексу (НДВІ) за 2021 – 2022 рр. для посівів олійних культур (ріпаку озимого та соняшнику), розташованих на Півдні України (у Херсонській, Одеській, Миколаївській, Запорізькій, Дніпропетровській, Кіровоградській областях). Дані супутникового моніторингу отримано на платформі GIMMS Global Agricultural Monitoring System, NASA. Використано знімки Terra MODIS NDVI у поєднанні з маскою посівів культур GDA Crop Specific Masks. Математичну обробку результатів і побудову моделей здійснювали за алгоритмами лінійної регресії, достовірність і точність моделей оцінювали за показниками F- та P-статистики. Побудову моделей, графічну роботу виконували в Python 3.13. **Результати.** Розроблено регресійні моделі врожаю ріпаку озимого та соняшнику за величиною НДВІ. Установлено, що вони мають високу адекватність (коефіцієнт детермінації для ріпаку озимого – 0,61, соняшнику – 0,70) і точність (середня абсолютна похибка для ріпаку озимого – 0,13 т/га, або 5,6%, для соняшнику – 0,16, або 7,8%.

Проте варто зазначити, що, попри високу точність, моделі мають лімітоване практичне значення, оскільки деталізована статистична оцінка підтверджує недостатній рівень математичної надійності моделі для ріпаку озимого (P -статистика для моделі ріпаку озимого — 0,13, соняшнику — 0,007 за мінімально необхідного рівня $P < 0,05$). Висновки. За параметрами сезонної динаміки величини супутникового НДВІ розроблено математичні моделі врожайності основних олійних культур, вирощуваних на Півдні України. Вони мають високу адекватність і точність, проте модель для ріпаку озимого має суто теоретичне значення через недостатній рівень надійності передбачень.

Ключові слова: нормалізований диференційний вегетаційний індекс, регресійний аналіз, ріпак озимий, соняшник, Степ України.

DOI: <https://doi.org/10.31073/agrovisnyk202503-07>

Дані аерокосмічного моніторингу є джерелом інформації для оцінювання стану посівів сільськогосподарських культур і прогнозування їх можливої продуктивності. На сучасному етапі розвитку аграрної науки тісна інтеграція агрономічних знань зі здобутками геоінформаційних систем та інформаційних технологій дає змогу виконувати побудову складних математичних моделей оцінювання росту і розвитку сільськогосподарських культур. В основу таких моделей зазвичай покладено вегетаційні індекси — показники, які розраховують за визначеними алгоритмами з використанням аерофотознімків у різних спектральних каналах. Нині найбільш широко використовуваним у науково-практичній діяльності є нормалізований диференційний вегетаційний індекс (НДВІ). Розрахунок його простий і полягає в обчисленні співвідношення між різницею та сумою спектральних каналів інфрачервоного (NIR) і червоного (Red) спектрів супутникових знімків [1]. Нормалізований диференційний вегетаційний індекс досить добре відображає загальний стан рослинності на території, дає можливість оцінити її життєздатність, ступінь розвитку, інтенсивність ростових процесів, а для культурних

рослин — спрогнозувати рівень урожайності [2]. Крім того, НДВІ є базисом для розрахунку комплексних вегетаційних індексів, скажімо, індексу стану рослинності (VCI) [3].

Перевагою НДВІ порівняно з іншими вегетаційними індексами є його широка доступність не лише для наукових спільнот, а й для агровиробників. Первинні аерофотознімки часто доступні на безоплатній основі, а їх обробка для розрахунку величини НДВІ не потребує великих затрат часу та обчислювальних потужностей [1, 4]. Більшість платформ агромоніторингу й комплексних експертних систем для відстеження стану посівів і коригування агротехнологій використовують саме цей вегетаційний індекс. Логічно, що основна увага науковців також приділяється саме моделям росту, розвитку та продуктивності сільськогосподарських культур, які ґрунтуються на НДВІ. Серед таких рішень перспективними і затребуваними практиками є моделі прогнозування можливої врожайності сільськогосподарських культур за даними супутникового моніторингу, зокрема величини НДВІ.

Південь України, який за природно-кліматичного районування належить до зони Степу, традиційно відомий як

зона вирощування олійних культур — ріпаку озимого та соняшнику. Щодо останньої культури, то Україна завжди була одним із найбільших експортерів насіння й соняшnikової олії в світовому масштабі [5]. Ріпак озимий — перспективна біоенергетична культура, площі вирощування якої на Півдні України поступово розширюються. З урахуванням стратегічного значення згаданих вище культур розроблення моделей оцінювання їх урожайності, за даними дистанційного моніторингу, є актуальним завданням сучасної аграрної науки, оскільки це питання ще недостатньо вивчене й висвітлене в Україні.

Мета досліджень — розробити математичні моделі врожайності основних олійних культур (ріпаку озимого та соняшнику), вирощуваних на Півдні України, з використанням даних аерокосмічного моніторингу їх посівів. Оцінити достовірність і точність моделей, визначити лімітувальні фактори та можливості в майбутньому науково-практичного використання розроблених моделей.

Матеріали та методи досліджень. Дослідження проведено з використанням даних середньомісячної величини НДВІ за 2021–2022 рр. для ріпаку озимого та соняшнику, розташованих на Півдні України (у Херсонській, Одеській, Миколаївській, Запорізькій, Дніпропетровській, Кіровоградській областях). Дані супутникового моніторингу отримано на платформі GIMMS (Global Inventory Modeling and Mapping Studies) Global Agricultural Monitoring System, NASA. Використано знімки Terra MODIS NDVI у поєднанні з маскою посівів культур GDA Crop Specific Masks. Супутникові знімки мають просторову роздільну здатність 250 м, якої недостатньо для деталізованого вивчення та моніторингу посівів на окремих полях і ділянках, але є придатними для виконання регіональних і зональних досліджень, що підтверджено в наукових роботах закордонних авторів,

скажімо, з вивчення змін у стані флори Австралії на тлі глобального потепління 1982–2006 рр. [6]. Із застосуванням маски посівів було вилучено всі інші рослинні об'єкти та угруповання, оцінювали регіональні величини вегетаційного індексу ексклюзивно для ріпаку озимого та соняшнику, що знижувало можливу похибку моделювання. Для ріпаку озимого враховували дані НДВІ за квітень — липень, для соняшнику — за квітень — вересень. Такі часові періоди зумовлені неоднорідністю застосовуваних агротехнологій у господарствах Півдня України та наявністю проміжних (післяжнивних і поукісних) посівів соняшнику на зрошуваних землях. Варто враховувати пересівання та підсівання культур, строки сівби, які не відповідають традиційним технологіям, а також різницю в тривалості періоду вегетації культур залежно від групи їх стиглості, застосування зрошення та інтенсивного мінерального удобрення, що зумовлює відхилення термінів збирання врожаю порівняно з традиційними технологіями. Дані щодо врожайності ріпаку озимого та соняшнику за роками досліджень у кожному регіоні одержано з офіційних статистичних джерел, таких, як «Статистичний щорічник України», у який за 2022 р. не було внесено даних по Херсонській і Запорізькій областях через бойові дії на цих територіях та часткову окупацію.

Математичну обробку результатів та побудову моделей здійснювали за алгоритмами лінійної регресії (метод звичайних найменших квадратів); достовірність і точність моделей оцінювали за показниками F-статистики та R-статистики для кожної вибірки [7]. Вибір лінійних алгоритмів пояснюється відносно малими розмірами вхідних вибірок даних згідно з центральною граничною теоремою [7], тому використання нелінійних моделей є недоцільним через високі ризики надмірного підгону кривої моделі та її наступної невідповідності за використання на сторонніх

1. Регресійна модель та статистичні показники врожайності ріпаку озимого на Півдні України за величиною супутникового НДВІ

Параметр (показник)	Значення
Модель	Урожайність = $1,3859 + 2,2263 \cdot \text{НДВІ}_{\text{квітень}} + 9,4153 \cdot \text{НДВІ}_{\text{травень}} - 13,0543 \cdot \text{НДВІ}_{\text{червень}} + 6,2780 \cdot \text{НДВІ}_{\text{липень}}$
Коефіцієнт детермінації	0,614
Коригований коефіцієнт детермінації	0,305
F-статистика	1,987
P-статистика	0,235
Похибка: абсолютна	0,131 т/га
відносна	5,6%

масивах даних [8]. Адекватність та якість підгону кривої моделей визначали за величиною коефіцієнта детермінації. Додатково було розраховано середні похибки (в абсолютній і відносній величинах). Розрахунки, побудову й тестування моделей, а також графічну роботу виконували в Python 3.13 [9–11].

Результати досліджень. Регресійна модель урожайності ріпаку озимого дає змогу оцінити продуктивність культури на основі змін індексу NDVI впродовж квітня — серпня (табл. 1). Модель описує вплив величини NDVI кожного місяця на врожайність (т/га). Установлено, що максимальний негативний вплив агроекологічних умов на ріпак (відповідно до величин регресійних коефіцієнтів моделі) спостерігався в червні.

Коефіцієнт детермінації (R^2) становить 0,614, що свідчить про взаємозв'язки між рівнем урожайності та показником НДВІ. Коригований коефіцієнт детермінації є нижчим (0,305), оскільки враховує кількість змінних у моделі. З огляду на невелику вхідну вибірку даних цей показник є достатнім. Значення F-статистики (1,987) та P-статистики (0,235) свідчать про статистичну значущість моделі, хоча P-значення не досягає її традиційного рівня ($< 0,05$). Абсолютна похибка становить 0,131 т/га, відносна — 4,1%, що підтверджує точне прогнозування врожайності культури.

Регресійна модель продуктивності соняшнику свідчить про те, що максимальний негативний вплив агро-екологічних умов на рослини культури

2. Регресійна модель та статистичні показники врожайності соняшнику на Півдні України за величиною супутникового НДВІ

Параметр (показник)	Значення
Модель	Урожайність = $1,5737 - 1,3924 \cdot \text{НДВІ}_{\text{квітень}} + 3,0640 \cdot \text{НДВІ}_{\text{травень}} - 3,2991 \cdot \text{НДВІ}_{\text{червень}} + 3,6228 \cdot \text{НДВІ}_{\text{липень}} - 4,8167 \cdot \text{НДВІ}_{\text{серпень}} + 4,5253 \cdot \text{НДВІ}_{\text{вересень}}$
Коефіцієнт детермінації	0,702
Коригований коефіцієнт детермінації	0,565
F-статистика	5,115
P-статистика	0,007
Похибка: абсолютна	0,163 т/га
відносна	7,8%

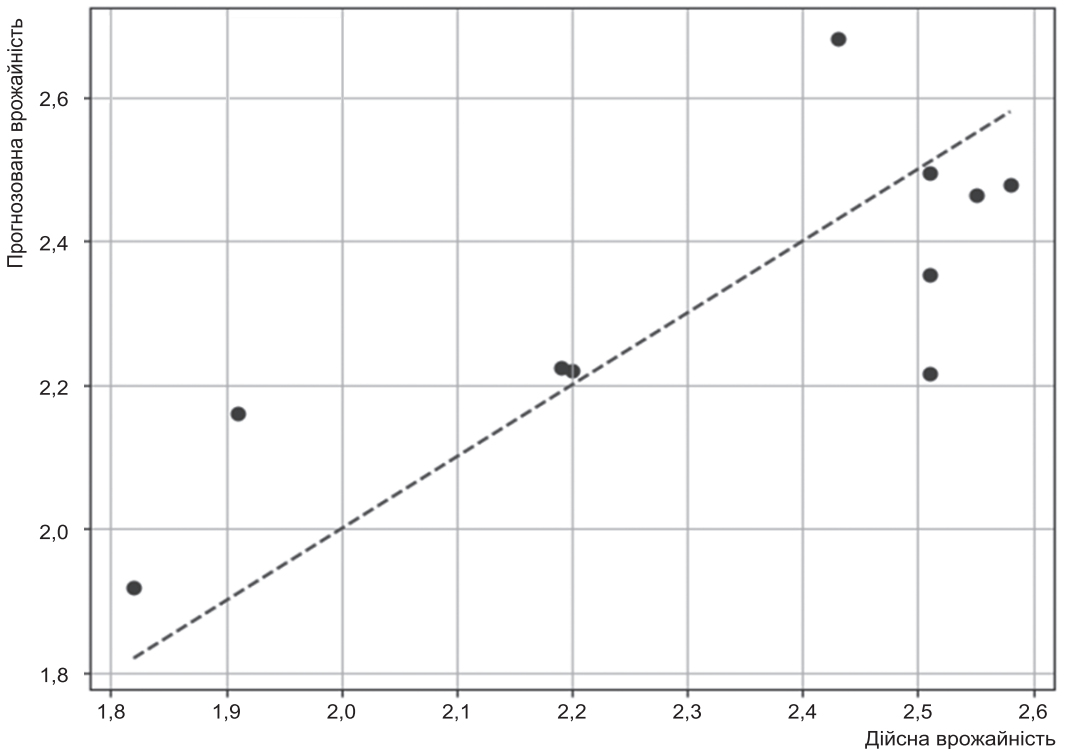


Рис. 1. Графічна апроксимація регресійної моделі врожайності ріпаку озимого залежно від величини НДВІ: ● — урожайність, т/га; --- — лінія тренду (для рис. 1, 2)

спостерігається в червні та серпні, які традиційно є досить посушливими в зоні Степу України (табл. 2). Коефіцієнт детермінації (R^2) дорівнює 0,702, що визначає достатній рівень пояснення змін у врожайності цією моделлю. Коригований коефіцієнт детермінації становить 0,565 з урахуванням кількості параметрів моделі. F-статистика (5,115) і P-статистика (0,007) підтверджують статистичну значущість моделі. Абсолютна похибка — 0,163 т/га, відносна — 7,8%, що є прийнятним для прогнозування.

Графічну апроксимацію розроблених математичних моделей наведено на рис. 1, 2. З урахуванням невеликих розмірів вибірок моделі мають лінійний характер для уникнення надмірного підгону кривих.

Варто зазначити, що подібні моделі фактично не мають аналогів в Україні.

Попередньо розроблені математичні моделі продуктивності олійних культур використовують подібні методичні підходи, але більшість із них або мають локальний характер (збір даних проведено в межах окремих дослідних ділянок чи полів), або не враховують маски посівів культур на регіональному рівні й необхідність виокремлення площ під орними землями, що було здійснено нами завдяки комбінації даних із Terra MODIS NDVI та GDA Crop Specific Masks [12]. Скажімо, було розроблено моделі продуктивності ріпаку озимого, соняшнику та сої посівної з використанням часової серії MODIS Terrain NDVI за 2012–2019 рр. для Херсонської області. Вони вирізнялися меншою точністю, ніж запропоновані нами: середня відносна похибка прогнозу врожайності становила 25,2% для соняшнику та 18,3% — для ріпаку

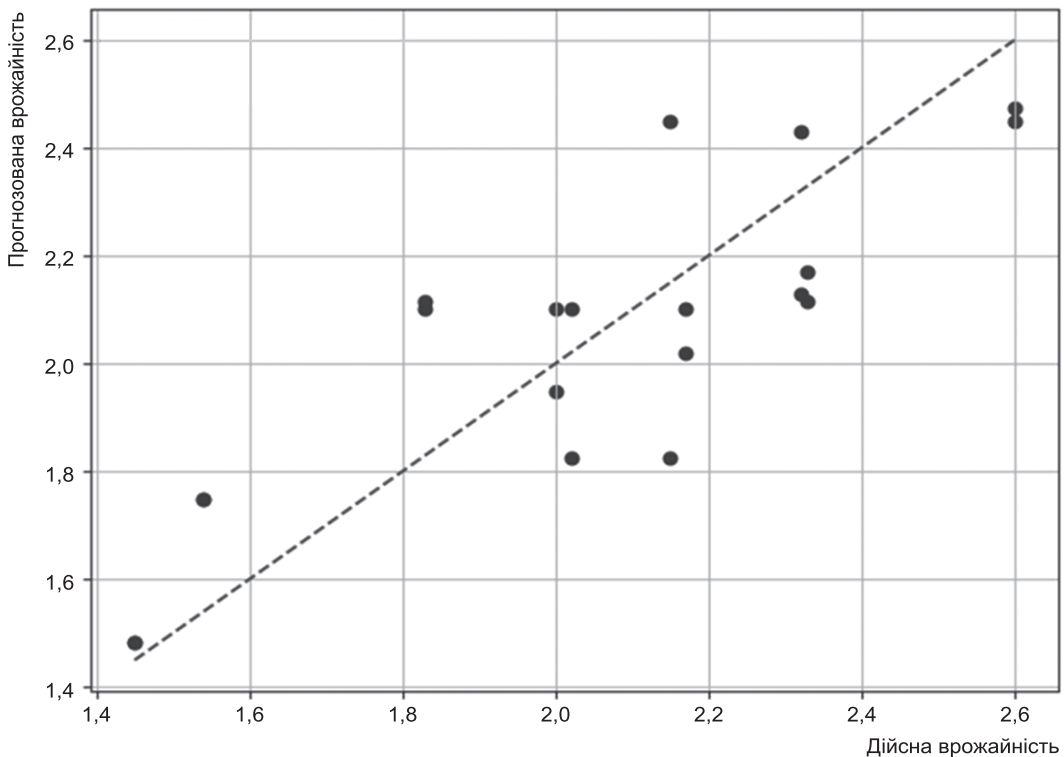


Рис. 2. Графічна апроксимація регресійної моделі врожайності соняшнику залежно від НДВІ

озимого [12]. Основним недоліком зазначеної роботи було використання загального, неспецифічного для кожної з культур значення НДВІ, оскільки маски посівів не застосовували.

Науково цінною є робота, виконана в умовах східного Китаю. Дослідниками встановлено, що найліпшу прогностичну точність та якість підгону кривої врожайності ріпаку озимого залежно від величини вегетаційного індексу забезпечувала модель ансамблевого навчання (коефіцієнт детермінації 0,925 за середньоквадратичної похибки 5,9%). Загалом розроблена нами модель на основі множинної лінійної регресії забезпечує порівняльний рівень точності за дещо нижчого рівня підгону кривої моделі. Слід зазначити, що перевагою нашого підходу є можливість одержання чіткого алгоритмічного рівняння для виконання розрахунків

і прогнозування врожайності культури, тоді як результати ансамблевого навчання не дають змоги отримати кінцеву конкретну математичну модель [13].

Варто зазначити, що представлена нами модель забезпечує високу якість і точність прогнозування врожайності соняшнику. Так, наприклад, розроблена за методом ансамблевого навчання модель продуктивності соняшнику в умовах південно-східної Угорщини забезпечила меншу точність (середньоквадратична похибка — 0,28–0,47 т/га) та нижчу якість підгону кривої моделей (коефіцієнт детермінації — 0,60) [14]. У роботі [15] моделюванням урожайності соняшнику за даними аерокосмічного моніторингу на регіональному рівні вдалося досягти коефіцієнта детермінації 0,730 за середньої відносної похибки в межах 4,2–7,2%. Ці

результати збігаються з даними, одержаними в наших дослідженнях. Проте варто відзначити вищу математичну обґрунтованість розробленої нами моделі порівняно з іншими моделями, оскільки нами додатково оцінено такі параметри достовірності, як F-статистика та P-статистика. Дослідження, виконані в умовах Туреччини в провінції Токат, результували за коефіцієнта детермінації моделей продуктивності соняшнику на рівні 0,74 та величини серед-

ньоквадратичної похибки 0,108 т/га, тобто ці показники збігалися з результатами, одержаними в нашій роботі (0,70; 0,163 т/га) [16].

Наступним етапом удосконалення моделей прогнозування врожаїв сільськогосподарських культур на основі даних дистанційного зондування Землі є залучення більших масивів вхідних даних і алгоритмів глибокого машинного навчання для отримання більш точних і надійних прогнозів [17].

Висновки

Уперше в Україні розроблено математичні моделі врожайності соняшнику та ріпаку озимого за величиною нормалізованого диференційного вегетаційного індексу для виконання прогнозування на регіональному рівні в зоні Степу України. Математичні моделі побудовано за алгоритмом множинної лінійної регресії та розроблено з урахуванням масок відповідних сільськогосподарських культур. Математико-статистичний аналіз підтвердив високу якість

і точність прогнозування (відносні похибки прогнозування становили 5,6% для ріпаку озимого та 7,8% — для соняшнику). Варто зазначити, що модель продуктивності ріпаку озимого має суто теоретичне значення через недостатній рівень надійності передбачень (P-статистика перевищує пороговий рівень 0,05). Надалі моделі можуть бути використані для науково-практичних та освітніх цілей, удосконалені за рахунок розширення вхідного набору даних.

Vozhehova R.¹, Lykhovyd P.², Hranovska L.³
Institute of Climate-Focused Agriculture of NAAS, 24 Maiatska Doroha Str., vil. Khlibodarske, Odessa district, Odessa oblast, 67667, Ukraine, e-mail: ¹izz.biblio@ukr.net, ²pavel.likhovid@gmail.com, ³G_Ludmila15@ukr.net; ORCID: ¹0000-0002-3895-5633, ²0000-0002-0314-7644, ³0000-0001-7021-3093

Mathematical modeling of oilseed yield in the South of Ukraine according to aerospace monitoring data

Goal. To develop mathematical models of the yield of the main oilseeds grown in the South of Ukraine, using data from aerospace monitoring of their crops. To evaluate the reliability and accuracy of models, determine the limiting factors and the possibilities of their scientific and practical use. **Methods.** The study was performed using data from the average monthly value of the normalized differential vegetation index (NDVI) for 2021–2022 for oilseeds (winter rape and sunflower) located in the South of Ukraine (in Kherson, Odesa, Mykolaiv, Zaporizhzhia,

Dnipropetrovsk, and Kirovograd regions). Satellite monitoring data was obtained on the GIMMS Global Agricultural Monitoring System platform, NASA. Images of Terra MODIS NDVI in combination with GDA Crop Specific Masks were used. Mathematical processing of the results and the construction of models were carried out using linear regression algorithms, and the reliability and accuracy of the models were evaluated using F- and P-statistics. Building models and graphic work were done in Python 3.13 environment.

Results. Regression models were developed of winter rape and sunflower yields by NDVI value. It was established that they had high adequacy (coefficient of determination for winter rape — 0.61, sunflower — 0.70), and accuracy (average absolute error for winter rape — 0.13 t/ha, or 5.6%, for sunflower — 0.16, or 7.8%). However, it should be noted that, despite the high accuracy, the models had limited practical value, since the detailed statistical evaluation confirmed the insufficient level of mathematical reliability of the model for winter rape (P-statistics for the model

of winter rape — 0.13, sunflower — 0.007 at the minimum required level $P < 0.05$). **Conclusions.** According to the parameters of the seasonal dynamics of the value of satellite NDVI, mathematical models of the yield of the main oilseeds grown in the South of Ukraine were developed. They had high adequacy and accuracy, but the

model for winter rape was of purely theoretical importance due to the insufficient level of reliability of predictions.

Key words: normalized differential vegetation index, regression analysis, winter rape, sunflower, Steppe of Ukraine.

DOI: <https://doi.org/10.31073/agrovysnyk202503-07>

Бібліографія

- Huang S., Tang L., Hupy J.P. et al. A commentary review on the use of normalized difference vegetation index (NDVI) in the era of popular remote sensing. *Journal of Forestry Research*. 2021. 32(1). P. 1–6. doi: 10.1007/s11676-020-01155-1
- Lopresti M.F., Di Bella C.M., Degioanni A.J. Relationship between MODIS-NDVI data and wheat yield: A case study in Northern Buenos Aires province, Argentina. *Information Processing in Agriculture*. 2015. 2(2). P. 73–84. doi: 10.1016/j.inpa.2015.06.001
- Yang Z., Di L., Yu G., Chen Z. Vegetation condition monitoring. *IEEE International Geoscience and Remote Sensing Symposium*. 2011. P. 3534–3537. doi: 10.1109/IGARSS.2011.6049984
- Jiang Z., Huete A.R., Chen J. et al. Analysis of NDVI and scaled difference vegetation index retrievals of vegetation fraction. *Remote Sensing of Environment*. 2006. 101(3). P. 366–378. doi: 10.1016/j.rse.2006.01.003
- Степенко С. Аналіз поточного стану аграрного сектору економіки України. *Наука Онлайн*. 2021. 12. 9 с.
- De Keersmaecker W., Lhermitte S., Hill M.J. et al. Assessment of regional vegetation response to climate anomalies: a case study for Australia using GIMMS NDVI time series between 1982 and 2006. *Remote Sensing*. 2017. 9(1). P. 34. doi: 10.3390/rs9010034
- Dodge Y. The concise encyclopedia of statistics. New York: Springer, 2008. 615 p.
- Bishop C.M., Nasrabadi N.M. Pattern recognition and machine learning. *J. of Electronic Imaging*. 2006. 4(4). 738 p. doi:10.1117/1.2819119
- Burton A.L. OLS (Linear) regression. *The Encyclopedia of Research Methods in Criminology and Criminal Justice*. 2021. 2. P. 509–514. doi: 10.1002/9781119111931.ch104
- Sureiman O., Mangera C.M. F-test of overall significance in regression analysis simplified. *Journal of the Practice of Cardiovascular Sciences*. 2020. 6(2). P. 116–122. doi: 10.4103/jpcs.jpcs_18_20
- Chai T., Draxler R.R. Root mean square error (RMSE) or mean absolute error (MAE). *Geoscientific Model Development Discussions*. 2014. 7(1). P. 1525–1534. doi: 10.5194/gmd-7-1247-2014
- Lykhovyd P. Forecasting oil crops yields on the regional scale using normalized difference vegetation index. *J. of Ecological Engineering*. 2021. 22(3). P. 53–57. doi: 10.12911/22998993/132436
- Hu H., Ren Y., Zhou H. et al. Oilseed rape yield prediction from UAVs using vegetation index and machine learning: A case study in East China. *Agriculture*. 2024. 14(8). P. 1317. doi: 10.3390/agriculture14081317
- Amankulova K., Farmonov N., Mukhtorov U., Mucsi L. Sunflower crop yield prediction by advanced statistical modeling using satellite-derived vegetation indices and crop phenology. *Geocarto International*. 2023. 38(1): 2197509. doi: 10.1080/10106049.2023.2197509
- Bognár P., Kern A., Pásztor S. et al. Testing the robust yield estimation method for winter wheat, corn, rapeseed, and sunflower with different vegetation indices and meteorological data. *Remote Sensing*. 2022. 14(12): 2860. doi: 10.3390/rs14122860
- Narin O.G., Abdikan S. Monitoring of phenological stage and yield estimation of sunflower plant using Sentinel-2 satellite images. *Geocarto International*. 37(5). P. 1378–1392. doi: 10.1080/10106049.2020.1765886
- Muruganatham P., Wibowo S., Grandhi S. et al. A systematic literature review on crop yield prediction with deep learning and remote sensing. *Remote Sensing*. 14(9): 1990. doi: 10.3390/rs14091990