

**МНОЖИННА РЕГРЕСІЯ ТА ПСЕВДООБЕРНЕННЯ В
ЗАДАЧАХ ДИСТАНЦІЙНОГО ЗОНДУВАННЯ ТА
ОЦІНЮВАННЯ ВМІСТУ ХЛОРОФІЛУ В РОСЛИННОСТІ**

В.О. ЯЦЕНКО, П.О. ХАНДРІГА, Д.О. ШВЕЦЬ

Розглянуто проблему використання спектрів відбиття рослинності для оцінювання вмісту хлорофілу. Описано новий метод оцінювання вмісту хлорофілу, який базується на псевдообернених матрицях та множинній регресії. Наведено результати порівняльного аналізу обчислення концентрації хлорофілу на основі множинної регресії та оптимізаційного методу.

ВСТУП

Проблема ризику екологічної обстановки, яка з часом ускладнюється, є однією з основних у програмі забезпечення безпеки існування людського суспільства [1, 2, 3]. У її розв'язанні значну роль відіграють системи моніторингу різних хімічних та біологічних компонентів навколишнього середовища з метою спостереження еволюції деградаційних процесів у природі, а також виявлення ексцесів, що виникають через аварійні ситуації техногенного походження або природні аномалії. До числа завдань моніторингу можна віднести: а) виявлення і прогнозування еволюції стану різних ділянок суходолу; б) раннє виявлення екологічних стресів; в) контроль стану техногенних об'єктів підвищеної небезпеки; г) контроль господарських соціально важливих об'єктів (наприклад, сільськогосподарські угіддя і ліси). Для розв'язання таких завдань найбільш прийнятні дистанційні спостереження, які вимагають розробки нових методів одержання інформації. Потрібно знайти зручні для дистанційних вимірів параметри об'єктів і способи тематичного дешифрування результатів вимірів.

Рослини є необхідним компонентом живого світу, а фотосинтетичний апарат зелених листків може бути використаний як чутливий індикатор стресових ситуацій [4]. Останнє обумовлено тим, що стреси різного походження впливають, насамперед, на вміст хлорофілу в листках. Цей параметр може бути оцінений безконтактним дистанційним методом, заснованим на використанні характеристик відбиття листків в оптичному діапазоні. Існує велика кількість публікацій, які вказують на реакцію пігментного апарату рослин у відповідь на дію стресу [1–6]. Відомі також роботи, що вказують на зниження хлорофілу в рослинності при витоках газу на магістральних газопроводах [2], вірусних захворюваннях [3], дефіциті азотного харчування [4–6] і вологи [7–10].

Вже тривалий час розробляється методологічний підхід до визначення хлорофілу в листках на основі кореляційного зв'язку цього параметра з величиною відносини двох екстремумів у графіку 1-ї похідної [4, 11–21]. Показано, що такий підхід дозволяє одержати рівняння регресії для розрахунку вмісту хлорофілу [18, 30], і що ця формула високостійка до внесків відбиття від ґрунту. Однак працездатність методу істотно залежить від можливості розрізнення двох екстремумів у графіку 1-ї похідної. Крім того, для реалізації методу необхідно розробити спосіб автоматичного пошуку положення екстремумів [20].

Вказані недоліки стимулювали пошук альтернативних методів визначення концентрації хлорофілу з використанням спектральної кривої відбиття. Застосування інформаційно-статистичного методу [30] та псевдообернення [25, 26] дозволило розробити новий метод та побудувати рівняння множинної регресії для концентрації хлорофілу.

ЕКСПЕРИМЕНТАЛЬНІ ДАНІ

У цій роботі як експериментальні дані використані спектральні криві листків рослин і відповідні значення концентрації хлорофілу. Спектральні криві отримані за допомогою спектрофотометра СФ-10, обладнаного інтегруючою сферою та системою реєстрації в цифровому вигляді. Концентрація хлорофілу визначалася хімічним методом Арнона [22]. На рис. 1 наведено приклад спектральних кривих для озимої пшениці при 100% проективного покриття. Для експериментів використані дані, отримані групою проф. С.М. Кочубей [4,16–21].

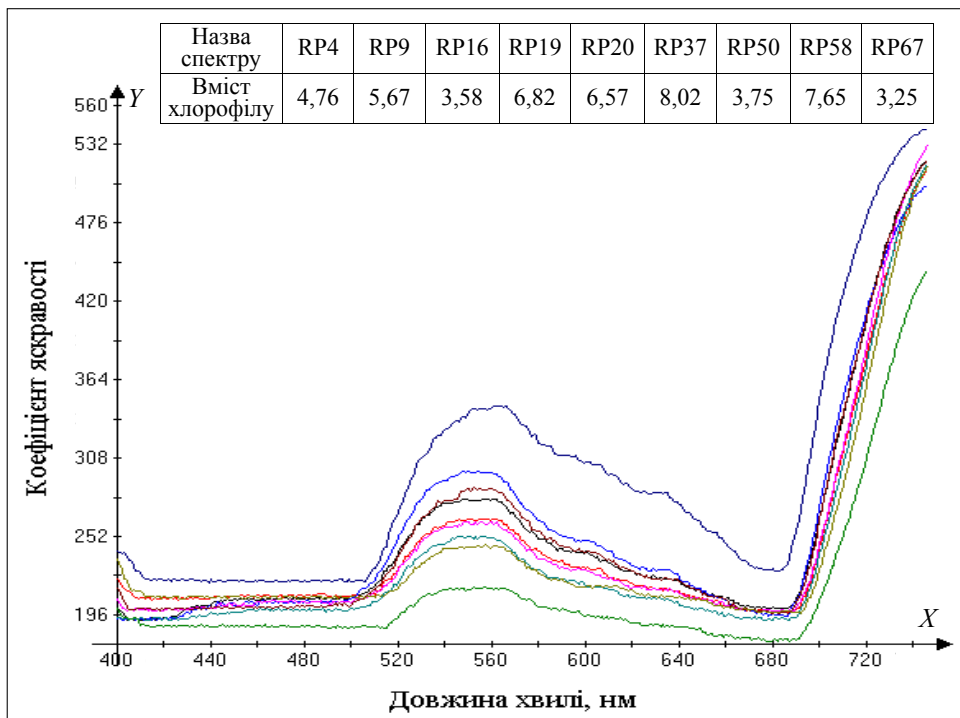


Рис. 1. Спектральні криві озимої пшениці при 100%-му проективному покритті

Варіації вмісту хлорофілу обумовлені розходженнями мінерального харчування і віком рослин. Система «рослинисть–грунт», яка була використана для з'ясування стійкості алгоритмів оцінки хлорофілу до впливу спектру відбиття від ґрунту при неповному проєктивному покритті, імітувалася за допомогою фізичних і графічних моделей. Перші здійснювалися шляхом покриття листками тільки певної частини поверхні вимірювальної кювети. Фоном служив чорнозем або пісок. У графічних моделях неповне проєктивне покриття моделювалося математичним додаванням спектральних кривих відбиття листків і фону із заданими ваговими коефіцієнтами. Для перевірки нормальності функції розподілу концентрації хлорофілу використано статистичний тест Колмогорова–Смірнова [23]. Значення концентрації хлорофілу використовувалися на етапі «навчання» із застосуванням множинної регресії. Набір кривих, з якого були вилучені помилкові спостереження, налічував 23 криві.

МЕТОД

Для оцінки вмісту хлорофілу за спектральними даними розроблено метод опису набору спектральних кривих, які відповідають різній концентрації хлорофілу, з використанням системи лінійних рівнянь. Оцінка здійснюється на основі псевдообернення.

Характерною особливістю спектральної кривої є наявність двох локальних максимумів, що знаходяться в області червоного краю спектру вимірювання (680...750 нм). Довжини хвиль, які відповідають цим максимумам знаходяться в певній залежності від вмісту хлорофілу в даній рослині.

У деяких випадках локальні максимуми першої похідної можуть бути представлені неявно, тобто мати форму точок перегину. Тому для побудови вектора характеристичних ознак спектральної кривої доцільно брати по декілька характеристичних точок з області першого (696...710 нм) та другого (716...724 нм) локальних максимумів. На рис. 2 показано графік першої похідної спектральної кривої RP9 в області червоного краю спектру вимірювання, з якого видно розташування першого та другого локальних максимумів.

Припускаючи, що структура залежності значення хлорофілу від характеристичних точок першої похідної спектральної кривої є лінійною, побудуємо регресійну модель

$$z = a_1x_1 + a_2x_2 + \dots + a_nx_n + b_1y_1 + b_2y_2 + \dots + b_my_m + c, \quad (1)$$

де z — значення хлорофілу; x_1, \dots, x_n — ординати точок з області першого локального максимуму; y_1, \dots, y_m — ординати точок з області другого локального максимуму; $k = (a_1, \dots, a_n, b_1, \dots, b_m, c)^T$ — невідомий вектор коефіцієнтів регресійної моделі (1).

Для обчислення вмісту хлорофілу за вектором характеристичних ознак необхідно розв'язати задачу ідентифікації, тобто знайти вектор невідомих параметрів k . Візьмемо l спектральних кривих, для яких відомі відповідні значення хлорофілу, і запишемо систему лінійних алгебраїчних рівнянь відносно параметрів $a_1, \dots, a_n, b_1, \dots, b_m, c$

$$\begin{cases} z_1 = a_1x_{11} + a_2x_{12} + \dots + a_nx_{1n} + b_1y_{11} + b_2y_{12} + \dots + b_my_{1m} + c, \\ z_2 = a_1x_{21} + a_2x_{22} + \dots + a_nx_{2n} + b_1y_{21} + b_2y_{22} + \dots + b_my_{2m} + c, \\ \dots \\ z_l = a_1x_{l1} + a_2x_{l2} + \dots + a_nx_{ln} + b_1y_{l1} + b_2y_{l2} + \dots + b_my_{lm} + c, \end{cases} \quad (2)$$

де j -е рівняння ($j = \overline{1, l}$) відповідає j -й спектральній кривій. У матричному вигляді систему (2) запишемо як

$$Ak = b, \quad (3)$$

$$\text{де } A = \begin{pmatrix} x_{11} & x_{12} & \dots & x_{1n} & y_{11} & y_{12} & \dots & y_{1m} & 1 \\ x_{21} & x_{22} & \dots & x_{2n} & y_{21} & y_{22} & \dots & y_{2m} & 1 \\ \dots & \dots & \dots & \dots & \dots & \dots & \dots & \dots & \dots \\ x_{l1} & x_{l2} & \dots & x_{ln} & y_{l1} & y_{l2} & \dots & y_{lm} & 1 \end{pmatrix}, \quad k = \begin{pmatrix} a_1 \\ a_2 \\ \dots \\ a_n \\ b_1 \\ b_2 \\ \dots \\ b_m \\ c \end{pmatrix}, \quad b = \begin{pmatrix} z_1 \\ z_2 \\ \dots \\ z_l \end{pmatrix},$$

$$A \in R^{l \times (n+m+1)}, \quad k \in R^{n+m+1}, \quad b \in R^l.$$

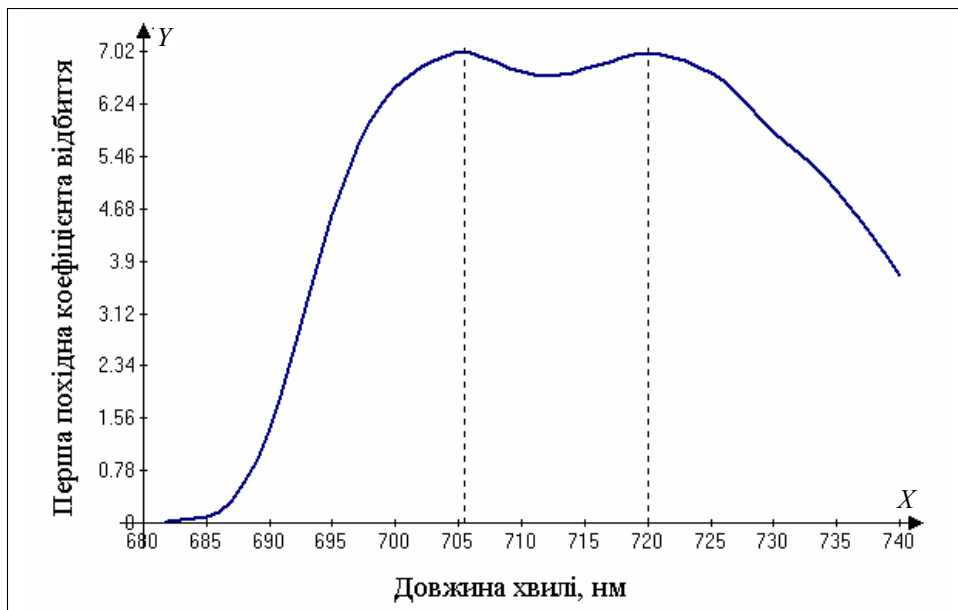


Рис. 2. Графік першої похідної спектральної кривої RP9 в області червоного краю спектру вимірювання

У загальному випадку матриця A має прямокутну форму, тому для знаходження вектора параметрів k використаємо теорію псевдообернених матриць [24]. Це дає змогу уникнути обмежень на розмірність матриці A та отримати не точкову оцінку вектора k , а деяку множину оцінок [24, 25]

$$\Omega(k) = \left\{ k : k = A^+b + Z(A)v, \forall v \in R^{n+m+1} \right\}.$$

Кожен елемент побудованої таким чином множини $\Omega(k)$ мінімізує нев'язку рівняння (3)

$$\|Ak - b\| \rightarrow \min_k.$$

Якщо вектор v покласти рівним нулю, то отримаємо розв'язок з найменшою нормою.

Запропонована модель дає можливість змінювати розмірність вектора характеристичних ознак, що впливає на точність визначення хлорофілу. Значним чином на якість ідентифікації впливає також параметр моделі l . Зокрема, при більшій кількості початкових (навчаючих) кривих розв'язок буде точнішим.

Цікавим та важливим з практичної точки зору є використання теорії збурень псевдообернених матриць. При надходженні до моделі нових спектральних кривих необхідно виконати лише корегування вектора параметрів, використовуючи формули Гревеля [24, 26, 28], а не обчислювати псевдообернену матрицю спочатку.

Розроблений метод реалізовано у вигляді прикладного програмного забезпечення (ППЗ) для двох варіантів використання (рис. 3). Перший, користувачський, призначений для експрес-оцінювання хлорофілу в рослинності і застосовується в польових умовах. У функції ППЗ WINCHL входить: одержання спектральних даних із спектрометра, обчислення концентрації хлорофілу і збереження результатів у базі даних.

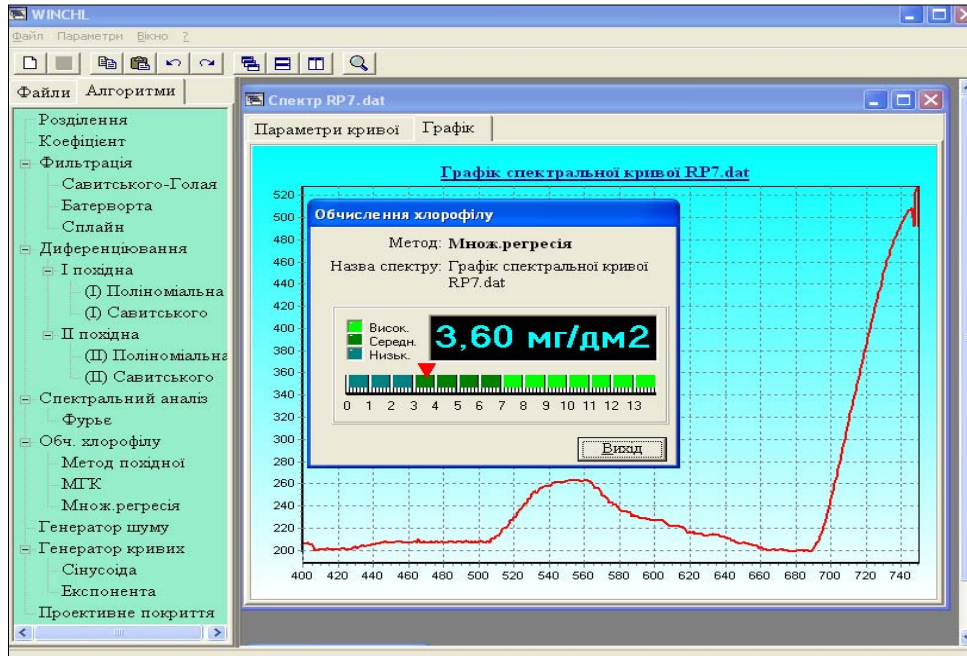


Рис. 3. Графічний інтерфейс програмного забезпечення для обчислення вмісту хлорофілу на основі множинної регресії

Розроблений варіант ППЗ простий у використанні, призначений для накопичення спектральних даних у базі даних і одержання швидких оцінок

вмісту хлорофілу в рослинності. Другий, дослідницький, розроблено для тестування широкого набору інструментів для обробки спектральних кривих і дослідження їхніх властивостей. Бібліотека ППЗ WINCHL містить алгоритми фільтрації і згладжування, Фур'є-аналізу, диференціювання, генетичного пошуку екстремуму, обчислення головних компонент [27], множинної регресії та генератор шумів, а також візуальний конструктор алгоритмів, що дозволяє дослідникові задати послідовність обробки спектральних кривих і їхні параметри на вибір та формувати різні алгоритми оцінювання вмісту хлорофілу. Дослідницький варіант дозволяє також моделювати шум і підбирати оптимальні параметри для їх фільтрації. Результати оцінювання вмісту хлорофілу, а також параметри алгоритму, який застосовується для оцінювання, зберігаються у базі даних, що дозволяє накопичувати статистику для різних алгоритмів.

Проведено комп'ютерне моделювання для навчаючих наборів з 6, 9, 12 та 23 кривих. З урахуванням результату обчислення коефіцієнтів k множинної регресії для 6 кривих рівняння (1) має вигляд

$$z = 2,95x_1 - 1,73x_2 - 0,87x_3 + 1,57x_4 - 2,15x_5 + 0,81x_6 - 2,95x_7 + 1,01x_8 + 0,83y_1 - 0,20y_2 - 1,15y_3 - 0,31y_4 + 3,03y_5 + 3,26. \quad (4)$$

ПОРІВНЯЛЬНИЙ АНАЛІЗ АЛГОРИТМІВ ВИЗНАЧЕННЯ ВМІСТУ ХЛОРОФІЛУ

Нами проведено порівняння методу множинної регресії з оптимізаційним підходом. Використано такі обмеження по довжині хвиль: для першого максимуму — (696...706), для другого — (717...724). Обмеження вибиралися таким чином, щоб пошук локальних екстремумів проводився у найбільш інформативних областях спектральної кривої. Для перевірки оптимізаційного алгоритму і візуалізації результатів створено програму, яка дає можливість увести похідні спектральних кривих з текстових файлів, відобразити їх графічно і знаходити локальні максимуми. Передбачена також можливість роздруковувати отримані результати. Для знаходження першого і другого локальних максимумів задаються обмеження на їхній пошук. Ці обмеження також можна задавати в інтерактивному режимі.

Після проведення чисельного експерименту була складена порівняльна таблиця (табл. 1). Пошук локальних максимумів проводився за допомогою методу граничного оцінювання (МГО), емпіричного методу (ЕМ) і генетичного алгоритму (ГА). За результатами проведеного порівняльного аналізу зроблено висновок про те, що МГО та ГА дають досить близькі результати.

При розробці оптимізаційного алгоритму для визначення екстремумів першої похідної спектральної кривої використовувалися дані, отримані в лабораторних умовах. Це мало певний вплив на форму спектральної кривої, зокрема, на більшості кривих були присутні два чіткі максимуми в області червоного краю (680...750 нм). Однак в реальних умовах експлуатації приладу спектральні криві будуть з деякими похибками. Причиною їх може бути, наприклад, певний відсоток проективного покриття, яке ввійшло до «знімку» рослинності. Тому важливим є тестування розробленого алгоритму на спектральних кривих з урахуванням проективного покриття. У табл. 2 наведено

порівняльні результати визначення екстремумів спектральних кривих без проєктивного покриття та з врахуванням 50% проєктивного покриття.

Таблиця 1. Порівняльна таблиця результатів

Файл	ЕМ		МПО		ГА	
	Довжина хвилі 1	Довжина хвилі 2	Довжина хвилі 1	Довжина хвилі 2	Довжина хвилі 1	Довжина хвилі 2
RP4	703	721	704	717	701	717
RP9	705	721	704	718	704	718
RP16	701	718	701	716	701	717
RP19	702	721	705	719	705	719
RP20	706	721	706	720	706	720
RP37	704	724	706	723	706	723
RP50	702	719	702	719	702	719
RP58	705	724	705	721	705	721
RP67	697	717	697	717	697	717

Як видно з табл. 2, розбіжності між знайденими максимумами становлять не більше 2 нм. Це свідчить про те, що розроблений алгоритм є інваріантним відносно певного рівня похибок, які з'являються при експлуатації приладу, і придатний для використання в різних умовах.

Таблиця 2. Локальні екстремуми спектральних кривих

Файл	Без проєктивного покриття		З 50% проєктивного покриття	
	Максимум			
	1-й	2-й	1-й	2-й
RP13	700	716	701	716
RP24	705	716	706	716
RP3	703	717	703	717
RP35	709	722	709	722
RP58	704	723	704	721
RP65	701	716	701	716

Для знаходження двох локальних максимумів першої похідної спектральної кривої використано генетичний алгоритм. Похибка у знаходженні максимальних значень по зведеній таблиці згаданих алгоритмів не перевищувала 2 нм. Генетичний алгоритм для спрощення реалізації та перевірки його ефективності був

реалізований в середовищі MatLab.

Порівняння методу множинної регресії та оптимізаційного підходу показало, що вони дають близькі результати за точністю обчислення вмісту хлорофілу. Метод множинної регресії виявився більш точним при обчисленні хлорофілу для 100% проєктивного покриття. У випадках, коли використовувалися рослини різних типів або проєктивне покриття було меншим, ніж 100%, оптимізаційний метод давав меншу похибку.

ВИСНОВКИ

1. Запропоновано новий метод оцінювання концентрації хлорофілу в рослинності, який базується на аналізі форми спектральної кривої в області червоного краю спектру (680...750 нм). Він заснований на використанні мо-

делі множинної регресії і псевдооберненої матриці. Розроблено програмне забезпечення, яке реалізоване у вигляді набору окремих модулів [29]. Детальний аналіз показав, що значних обчислювальних витрат алгоритм оцінки вмісту хлорофілу не вимагає. У перспективі це дозволяє використовувати метод множинної регресії разом з методом головних компонентів для більш точної оцінки концентрації на основі вбудованих і мобільних обчислювальних систем.

2. Метод визначення вмісту хлорофілу показав досить високі по точності результати. Побудова моделі множинної регресії та застосування теорії псевдообернення матриць дозволило значно підвищити точність даного методу та забезпечити інваріантність відносно похибок вхідних даних. Можливість оперативно задавати параметри моделі дозволяє автоматично, з необхідною точністю, визначати вміст хлорофілу в сільськогосподарській культурі, що, в свою чергу, покращує прогнозування врожайності.

Роботу виконано в рамках гранту УНТЦ № 2614.

ЛІТЕРАТУРА

1. *Davids C., Tyler A.N.* Detecting contamination-induced tree stress within the Chernobyl exclusion zone // *Remote Sens. Environ.* — 2003. — **85**. — P. 30–38.
2. *Smith K.L., Steven M.D., Coll J.J.* Use of hiper-spectral derivative ratios in the red-edge region to identify plant stress responses to gas leaks // *Remote Sens. Environ.* — 2004. — **92**. — P. 207–217.
3. *Changes in reflectance spectrum characteristics of Nicotiana debney plant under the influence of viral infection / V.P. Polischuk, T.M. Shadchina, T.I. Kompanetz et al.* // *Arh. Phytopath. Pflanz.* — 1997. — **31**. — P. 115–119.
4. *Кочубей С.М., Кобец Н.И., Шадчина Т.М.* Спектральные свойства растений как основа методов дистанционной диагностики. — Киев: Наук. думка, 1990. — 136 с.
5. *Differentiation among effects of nitrogen fertilization treatments on conifer seedlings by foliar reflectance: a comparison of methods / J.A. Moran, A.K. Mitchell, G. Goodmanson, K.A. Stockburger* // *Tree Physiology.* — 2000. — **20**. — P. 1113–1120.
6. *Estimating leaf nitrogen concentration in ryegrass (Lolium spp.) pasture using the chlorophyll red-edge: theoretical modelling and experimental observations / D.W. Lamb, M. Steyn-Ross, P. Schaare et al.* // *Int. J. Remote Sens.* — 2002. — **23**. — P. 3619–3648.
7. *Riggs G.A., Running S.W.* Detection of canopy water stress in conifers using the airborne imaging spectrometer // *Remote Sens. Environ.* — 1991. — **35**. — P. 51–68.
8. *Detecting vegetation leaf water content using reflectance in the optical domain / P. Ceccato, S. Flasse, S. Tarantola et al.* // *Remote Sens. Environ.* — 2001. — **77**. — P. 22–33.
9. *Ceccato P., Flasse S., Gregoire J.M.* Designing a spectral index to estimate vegetation water content from remote sensing data. 2. Validation and applications // *Remote Sens. Environ.* — 2002. — **82**. — P. 198–207.
10. *Bowyer P., Danson F.M.* Sensitivity of spectral reflectance to variation in live fuel moisture content at leaf and canopy level // *Remote Sens. Environ.* — 2004. — **92**. — P. 297–308.
11. *Horler D.N.H., Dokray M., Barber J.* The red edge of plant leaf reflectance // *Int. J. Remote Sens.* — 1983. — **4**. — P. 273–288.
12. *Pinar A., Curran P.J.* Grass chlorophyll and the reflectance red edge // *Int. J. Remote Sens.* — 1996. — **17**. — P. 351–357.

13. *Vegetation stress detection through chlorophyll a+b estimation and fluorescence effects on hyperspectral imagery* / P.J. Zarco-Tejada, J.R. Miller, G.H. Mohammed at al. // *J. Environ. Qual.* — 2002. — **31**. — P. 1433–1441.
14. *Steady-state chlorophyll a fluorescence detection from canopy derivative reflectance and double peak red edge effect* / P.J. Zarco-Tejada, J.C. Pushnik, S. Dobrowski, S.L. Ustin // *Remote Sens. Environ.* — 2003. — **84**. — P. 283–294.
15. *Chlorophyll fluorescence Detection with a high-spectral resolution spectrometer through in-filling of the O2-A band as function of water stress in olive trees* / P.J. Zarco-Tejada, O. Pérez-Priego, G. Sepulcre-Cantó at al. // *Remote Sensing of Vegetation Fluorescence*, 17–19 Nov. 2004, Montreal, Canada. — **90**. — P. 337–357.
16. Кочубей С.М., Кобец Н.И., Шадчина Т.М. Количественный анализ формы спектральной кривой отражения листьев растений как способ тестирования их состояния // *Физиология и биохимия культурных растений.* — 1988. — **20**. — P. 535–539.
17. Кочубей С.М. Сравнение информативных возможностей многозональной съемки и спектроскопии высокой разрешающей способности при дистанционном зондировании растительного покрова // *Космічна наука і технологія.* — 1999. — **5**, № 2/3. — С. 41–48.
18. Кочубей С.М. Оценка основных параметров сельскохозяйственных посевов по спектру отражения растительности в оптическом диапазоне // *Космічна наука і технологія.* — 2003. — **9**. — С. 185–190.
19. Kochubey S.M., Bidyuk P.I. A Novel Approach to Remote Sensing of Vegetation Conference «AeroSense. Technologies and Systems for Defence & Security», 21–25 April 2003, Orlando USA, Proc. of SPIE. — **5093**. — P. 327–329.
20. *Estimation of chlorophyll concentration in vegetation using global optimization approach* / V.A. Yatsenko, S.M. Kochubey, P.M. Pardalos, L. Zhan // *SPIE Conference «AeroSense. Technologies and Systems for Defence & Security»*, 21–25 April 2003, Orlando USA, Proc. of SPIE. — **5071**. — P. 50–59.
21. Кочубей С.М. Аппаратура и методы дистанционного зондирования растительности в оптическом диапазоне // *Космическая наука и технологии.* — 2002. — **8**. — С. 271–275.
22. Arnon D.I. Copper enzymes in isolated chloroplasts. Polyphenoloxidase in *Beta vulgaris* // *Plant Physiol.* — 1949. — **24**. — P. 1–15.
23. Орлов А.И. Прикладная статистика. Учебник. — М.: Экзамен, 2004. — 671 с.
24. Альберт А. Регрессия, псевдоинверсия, рекуррентное оценивание. — М.: Наука, 1977. — 305 с.
25. Кириченко Н.Ф., Лепеха Н.П. Псевдообращение в задачах управления и наблюдения // *Автоматика.* — 1993. — № 5. — С. 69–81.
26. Кириченко Н.Ф. Аналитическое представление возмущений псевдообратных матриц // *Кибернетика и системный анализ.* — 1997. — № 2. — С. 98–107.
27. Comon P. Independent component analysis. A new concept? // *Signal Processing*, 1994. — № 36(3). — P. 287–314.
28. Фаддеев Д.К., Фаддеева В.Н. Вычислительные методы линейной алгебры. — М.: Лань, 2002. — 736 с.
29. *Optical spectrometer and software for remote sensing of vegetation* / V.A. Yatsenko, V.V. Donets, S.M. Kochubey at al. // *Advanced Optoelectronics and Laser*, Yalta, Crimea, Ukraine, September 12–17, 2005, IEEE. — **1**. — P. 303–305.
30. Хандрига П.А., Яценко В.А. Информационно-статистический метод оценивания содержания биохимических компонентов в растительности // *Системні дослідження та інформаційні технології.* — 2006. — № 1. — С. 119–132.

Надійшла 27.01.2006