

ДВОКОМПОНЕНТНА АДАПТИВНА МОДЕЛЬ ДИНАМІКИ ВЕГЕТАЦІЙНИХ ІНДЕКСІВ ДЛЯ ПРОГНОЗУВАННЯ УРОЖАЙНОСТІ СІЛЬСЬКОГОСПОДАРСЬКИХ КУЛЬТУР

М. В. Мачуляк

Західноукраїнський національний університет
11, Львівська, Тернопіль, 46009, Україна
Email: Mvmach9@gmail.com

Запропоновано двокомпонентну адаптивну модель динаміки вегетаційних індексів, яка на відміну від відомих містить дискретну адаптивну компоненту на основі комбінації поточних та історичних даних та неперервну апроксимаційну компоненту на базі системи диференціальних рівнянь Моно. Розроблено метод ідентифікації дворівневої адаптивної моделі урожайності, який використовує пояснюючі змінні у вигляді моделей вегетаційних індексів NDVI та MTCI замість їх безпосередніх значень. Запропонована архітектура геоінформаційної системи на базі PostGIS забезпечує ефективну інтеграцію просторових даних від БПЛА та LiDAR систем. Математична модель виявлення ущільнених ділянок ґрунту на основі рівнянь Моно показала підвищення точності у півтора рази порівняно з лінійними методами. Експериментальна апробація на даних пшениці та рису підтвердила ефективність підходу з середньою похибкою прогнозування урожайності п'ятькома з трьома відсотка для ансамблевих методів. Отримані результати демонструють перспективність застосування запропонованого підходу для розвитку систем точного землеробства та підтримки прийняття агрономічних рішень.

Ключові слова: вегетаційні індекси, модель Моно, прогнозування урожайності, адаптивне моделювання, дистанційне зондування, геоінформаційні системи, точне землеробство, підтримка прийняття рішень, екологічний моніторинг

Вступ. Забезпечення продовольчої безпеки та підвищення ефективності агропромисловості є критично важливими завданнями сучасності. Традиційні методи оцінки урожайності базуються на статистичних даних регіонального або національного рівня, що не забезпечує необхідної просторової та часової роздільності для прийняття оперативних управлінських рішень [1]. Сучасні технології точного землеробства, зокрема використання безпілотних літальних апаратів (БПЛА) у поєднанні з геоінформаційними системами (ГІС), дозволяють отримувати високоточну інформацію про стан посівів через вегетаційні індекси. Нормалізований різницевий вегетаційний індекс (NDVI) та індекс хлорофілу наземної рослинності (MTCI) є надійними індикаторами фотосинтетично активної біомаси та вмісту хлорофілу відповідно [2]. Однак існуючі підходи до моделювання динаміки вегетаційних індексів мають суттєві обмеження: статистичні моделі не враховують нелінійну природу біологічних процесів, а моделі машинного навчання вимагають великих обсягів тренувальних даних [3].

Мета дослідження - розробка двокомпонентної адаптивної моделі динаміки вегетаційних індексів та методу ідентифікації дворівневої моделі урожайності для підвищення точності прогнозування в умовах обмежених даних.

Теоретичні основи дослідження. Сучасні підходи до моделювання урожайності можна класифікувати на три основні групи: статистичні, імітаційні та методи машинного навчання [5]. Статистичні моделі базуються на емпіричних залежностях між вегетаційними індексами та урожайністю. Baez-Gonzalez et al. [6] використовували множинну лінійну регресію для зв'язку NDVI з урожайністю кукурудзи, отримавши коефіцієнт детермінації $R^2 = 0.65$. Однак такі моделі не враховують динамічну природу росту рослин. Імітаційні моделі, такі як DSSAT та APSIM, моделюють біофізичні

процеси росту, але вимагають великої кількості параметрів і складні для калібрування [7]. Моделі машинного навчання демонструють високу точність, але мають обмежені пояснювальні властивості. Zhang et al. [8] використовували згорткові нейронні мережі для прогнозування урожайності пшениці з точністю $RMSE = 0.12$ т/га.

Перспективним напрямком є використання диференціальних рівнянь для моделювання біологічних процесів. Модель Моно, спочатку розроблена для опису росту мікроорганізмів, успішно адаптується для моделювання росту вищих рослин [9].

Постановка проблеми. Побудова ГІС на базі PostGIS є оптимальним рішенням для інтеграції просторових та атрибутивних даних. PostGIS забезпечує понад 400 просторових функцій, підтримку растрових даних та ефективну індексацію за допомогою GiST індексів. Структура основних відношень ГІС базується на центральній просторовій сутності – полі (farm_fields). Ключові компоненти включають:

- farm_fields: межі полів (field_id, name, area_ha, field_geometry);
- field_blocks: технологічні ділянки (block_id, field_id, block_geometry);
- yield_data: щорічні дані урожайності (yield_id, field_id, year, crop_type, yield_t_ha);
- vegetation_indices: динаміка вегетаційних індексів (veg_id, field_id, capture_date, ndvi_mean, source)

Для зберігання растрових даних використовується гібридний підхід: метадані зберігаються в реляційних таблицях (canopy_height_metadata, dtm), тоді як самі растрові файли розміщуються у файлової системі або хмарному сховищі. Така архітектура забезпечує ефективний доступ до просторових характеристик рослинного покриву за даними LiDAR.

Відомо, що ущільнення ґрунту призводить до зниження врожайності на 15–30 %, що обумовлює необхідність своєчасного виявлення ущільнених ділянок. Для цього у роботі запропоновано використовувати залежність між висотою рослин $X(t)$ та опором проникненню ґрунту, формалізовану на основі моделі Моно [4].

Аналіз експериментальних даних показує S-подібний характер залежності, який адекватно описується системою диференціальних рівнянь Моно:

$$\begin{cases} \frac{dX(t)}{dt} = p_1 \frac{X(t)S(t)}{p_2+S(t)} \\ \frac{dS(t)}{dt} = -p_3 \frac{X(t)S(t)}{p_2+S(t)} \end{cases} \quad (1)$$

де: $X(t)$ – висота рослин; $S(t)$ – потенціал росту; t – час; p_1, p_2, p_3 – параметри моделі, що визначають інтенсивність росту та споживання ресурсу.

Для зменшення впливу шумів, зумовлених похибками фотограмметрії та лазерного сканування, експериментальні дані попередньо згладжувалися гаусівським фільтром, який задається операцією згортки [6]:

$$Y[i] = \sum_{j=-k}^k D[i-j]K[j] \quad (2)$$

де: D – вихідні дані; K – ядро Гаусса; k – радіус ядра згладжування.

У випадку одновимірного згладжування ядро фільтра відповідає щільності нормального розподілу із нульовим математичним сподіванням:

$$G(x) = \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{x^2}{2\sigma^2}} \quad (3)$$

де: x – відстань від центру ядра; σ – стандартне відхилення, ступінь згладжування даних.

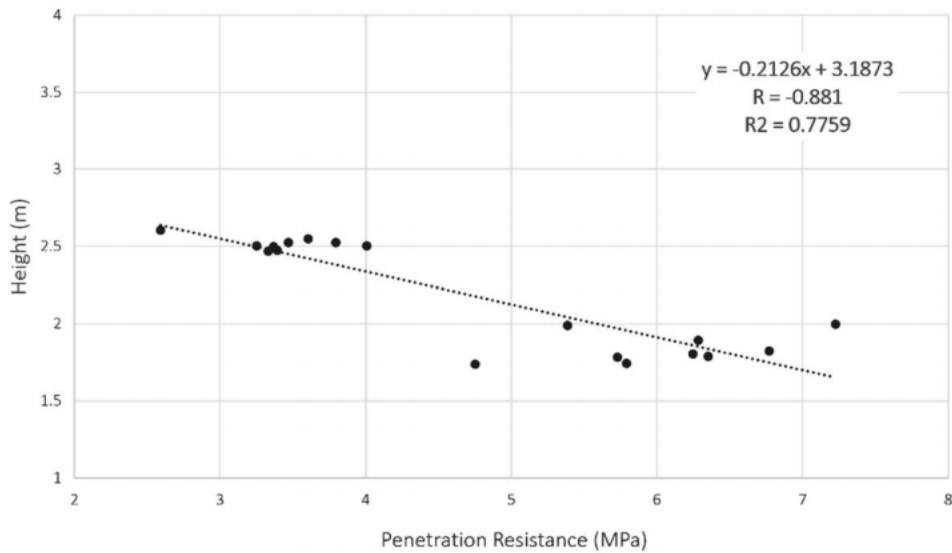


Рис. 1. Графік лінійної регресії між висотою та опором проникненню ґрунту

Після згладжування та нормалізації даних було виконано регресійний аналіз, який дозволяє розглядати локальну ділянку S-подібної кривої моделі Моно як квазілінійну залежність між висотою рослин і опором проникненню ґрунту. Результати такої апроксимації наведено на рисунку 1, де показано лінійну регресію між висотою рослин та опором проникненню ґрунту, що підтверджує негативний вплив ущільнення на параметри росту рослин.

Для підвищення достовірності подальшого моделювання було виконано попередню обробку експериментальних даних, яка включала видалення викидів, згладжування часових рядів гаусівським фільтром та нормалізацію висотних показників, отриманих за даними LiDAR. Результати зазначених етапів попередньої обробки наведено на рис. 2, де показано зменшення шумової складової та вирівнювання тренду висоти рослин.

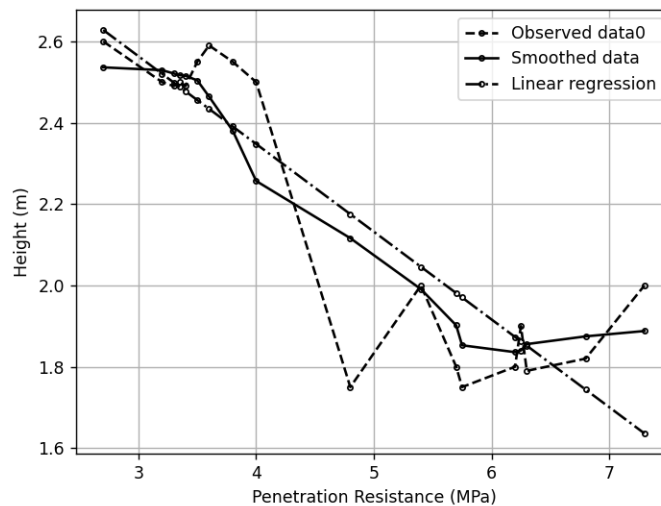


Рис. 2. Результати попередньої обробки даних

Для врахування реальних значень висоти та усунення впливу локального мікрорельєфу застосовувалося лінійне перетворення:

$$X = X^r - X^{min} \quad (4)$$

де: X^r – виміряна висота рослин за даними LiDAR; X^{min} – мінімальне значення висоти в межах досліджуваної ділянки; X – приведені значення висоти, що використовується в моделі.

Результати обговорення. Застосування перетворення (4) дозволило привести експериментальні дані до єдиної шкали та забезпечити коректну ідентифікацію параметрів моделі Моно. На основі нормалізованих даних було побудовано модель Моно співвідношення між ущільненістю ґрунту та висотою рослин, яка відображає нелінійний характер впливу ущільнення на ріст рослин.

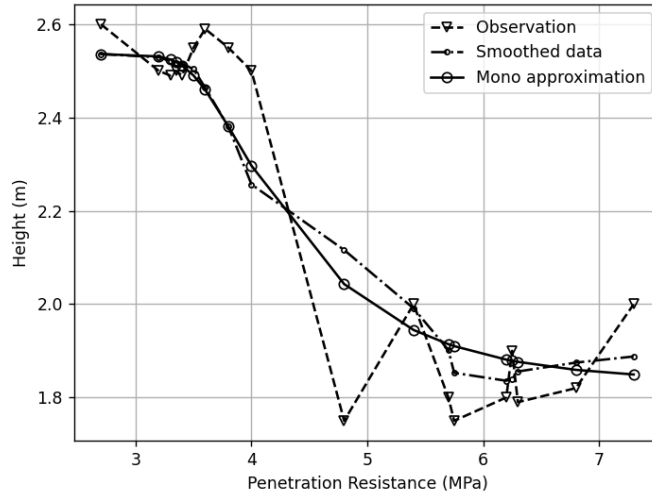


Рис. 3. Модель Моно співвідношення між ущільненістю ґрунту та висотою рослин

Порівняльний аналіз точності апроксимації показав, що модель Моно має вищу прогностичну здатність у порівнянні з лінійною регресією. Зокрема, для лінійної моделі максимальна відносна похибка становить 16,3 %, а середня — 4,5 %, тоді як для моделі Моно відповідні значення зменшуються до 10,4 % та 3,3 %. Це підтверджує доцільність використання нелінійної моделі Моно для аналізу впливу ущільнення ґрунту на параметри росту рослин. Для валідації запропонованого підходу використовувалися дані динаміки NDVI по 15 полях пшениці з північної Бельгії протягом трирічного періоду (2018-2020).

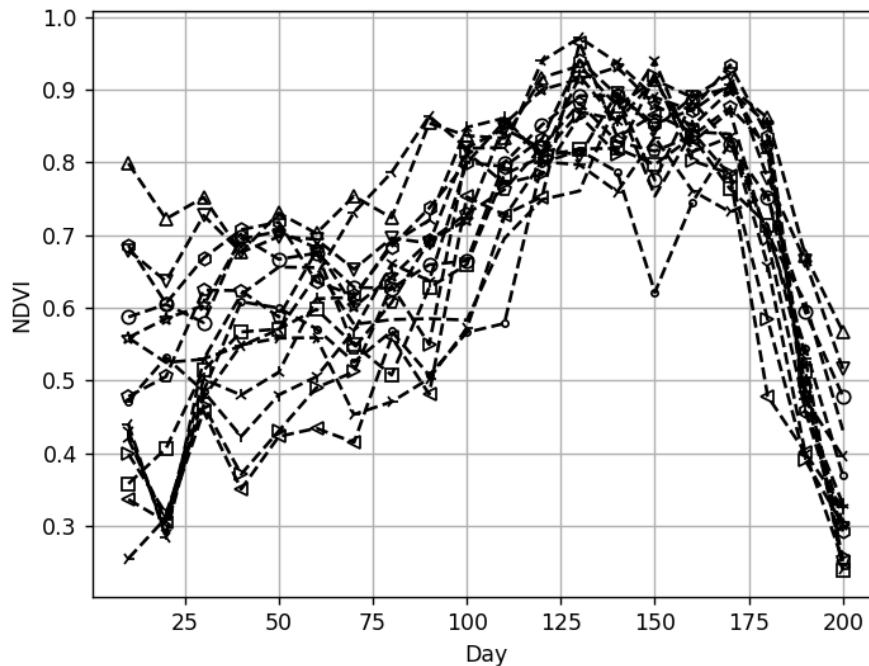


Рис. 4. Динаміка вегетаційного індексу NDVI

Перехід до кумулятивних значень NDVI забезпечив стабільність процесу моделювання та усунення випадкових коливань вихідних даних. Кумулятивні значення

формували щільний пучок траєкторій, що свідчить про узгодженість динаміки розвитку рослин на різних полях.

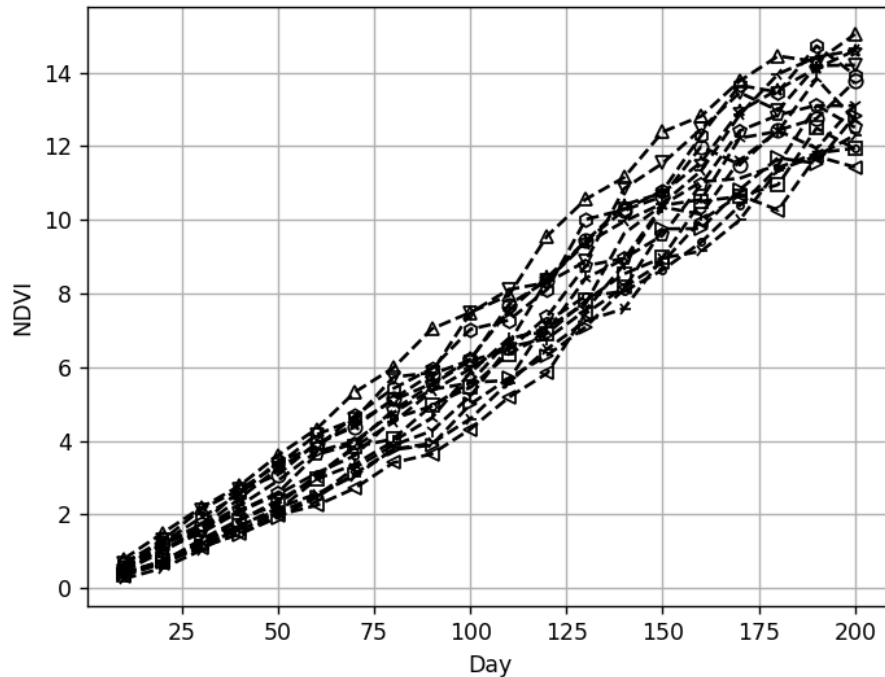


Рис. 5. Динаміка кумулятивних значень вегетаційного індексу

Апроксимація першої траєкторії за допомогою моделі Моно показала високу якість наближення з максимальною відносною похибкою 9,2% та середньою відносною похибкою 5,0%.

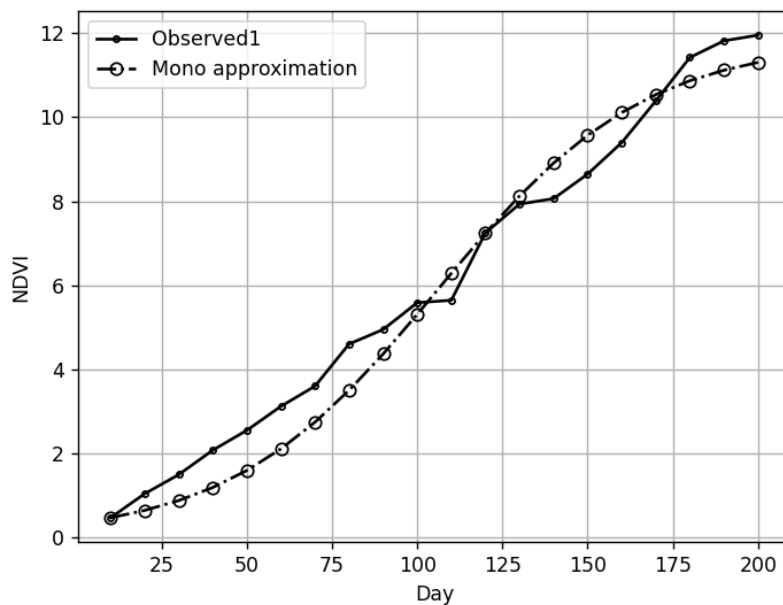


Рис.6. Апроксимація кумулятивних значень за допомогою моделі Моно

Ефективність адаптивного прогнозування була протестована на траєкторіях різної інтенсивності росту. Четверта траєкторія (нормальний ріст) показала максимальну похибку прогнозу 6,6% та середню 1,5%, а сьома траєкторія (знижена інтенсивність) – 5,2% та 1,4% відповідно. Графічна візуалізація результатів адаптивного прогнозування наведена на рис.7.

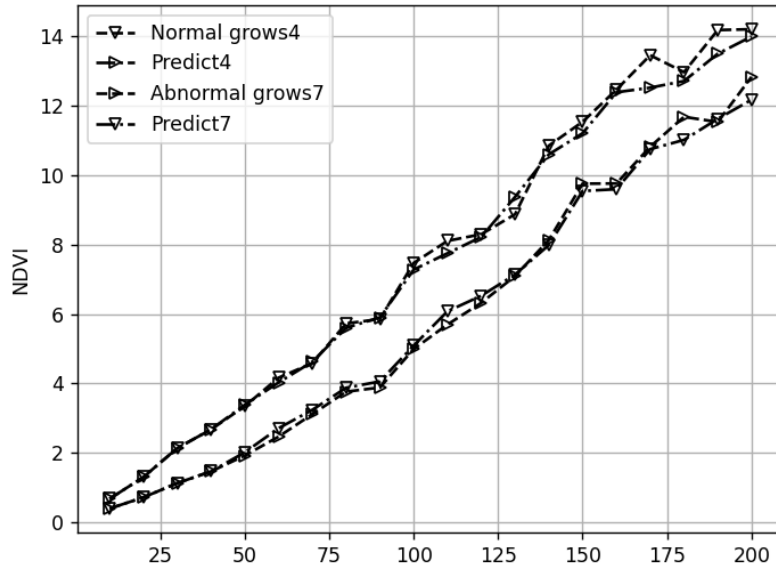


Рис. 7. Адаптивний прогноз для траєкторій різної інтенсивності

Для прогнозування урожайності пшениці були розроблені лінійна та нелінійна моделі на основі максимальних кумулятивних значень NDVI:

$$Y(f, Z) = L(\max_d(X_{NDVI}(f, Z, d))) \quad (5)$$

де: Y – урожайність; L – лінійний одновимірний оператор; X_{NDVI} – кумулятивне значення вегетаційного індексу NDVI; f – ідентифікатор поля; Z – сезон вирощування урожаю; d - день року.

Також побудуємо нелінійну залежність типу випадкового лісу

$$Y(f, Z) = RF(\max_d(X_{NDVI}(f, Z, d))) \quad (6)$$

де: RF – оператор побудови залежності типу випадкового лісу.

Для прогнозування урожайності пшениці були розроблені лінійна та нелінійна моделі на основі максимальних кумулятивних значень NDVI згідно з формулами (5) та (6). Результати застосування лінійної моделі, що використовувала 50% даних для навчання, демонструються на рис. 8.

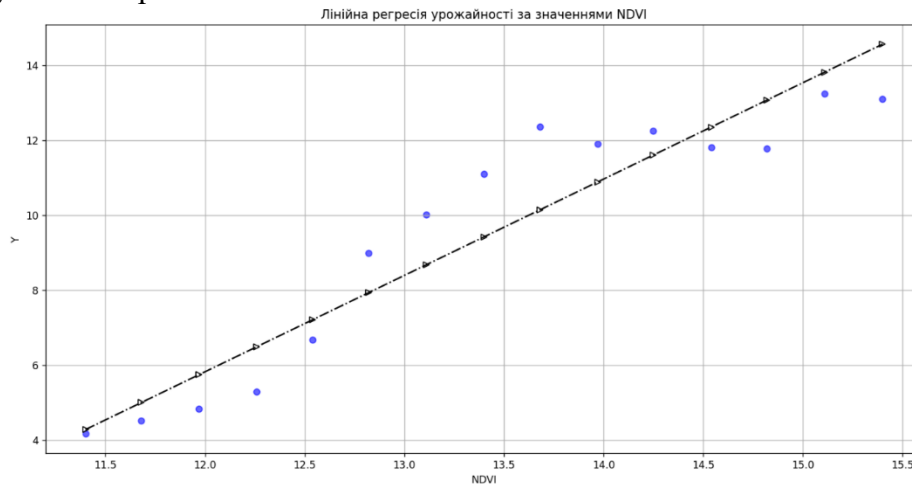


Рис. 8. Лінійна модель урожайності пшениці

Модель випадкового лісу з використанням 100 випадкових дерев показала дещо іншу структуру залежності, що відображено на рис. 9.

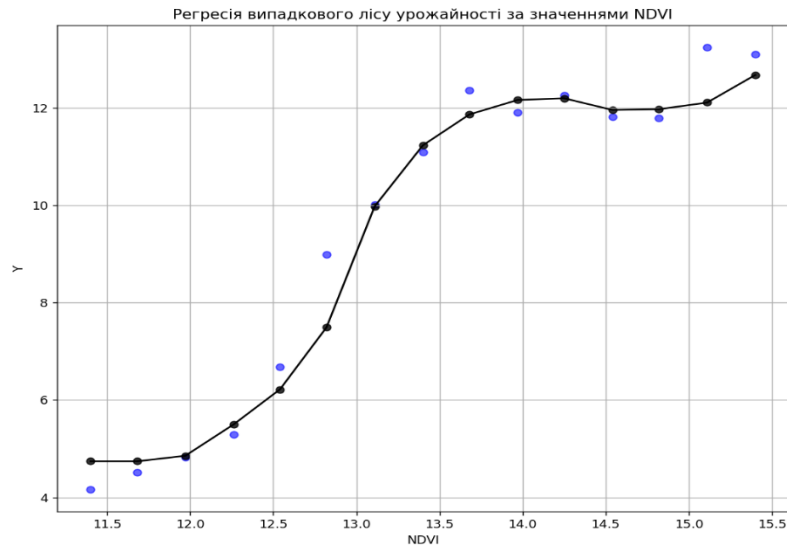


Рис. 9. Модель випадкового лісу урожайності пшениці

Результати моделювання урожайності пшениці: лінійна модель: максимальна похибка 8.9%, середня 4.8%; випадковий ліс: максимальна похибка 8.6%, середня 6.2%.

Дворівневий підхід передбачає окреме моделювання динаміки факторів урожайності (NDVI та МТСІ) з наступним прогнозуванням урожайності на основі модельованих значень індексів. Для експериментальної перевірки використовувався набір з 21 траєкторії динаміки вегетаційних індексів рису в нормальних та стресових умовах. Часові профілі спостережених траєкторій NDVI характеризуються значною варіабельністю, що показано на рис. 10.

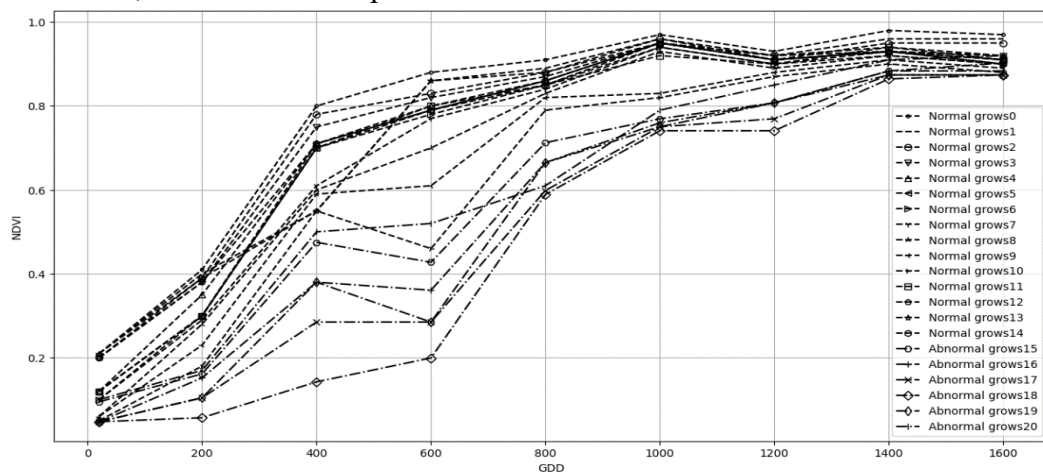


Рис. 10. Часові профілі NDVI

Аналогічно, траєкторії МТСІ демонструють складну динаміку з різними патернами розвитку, що відображено на рис. 11.

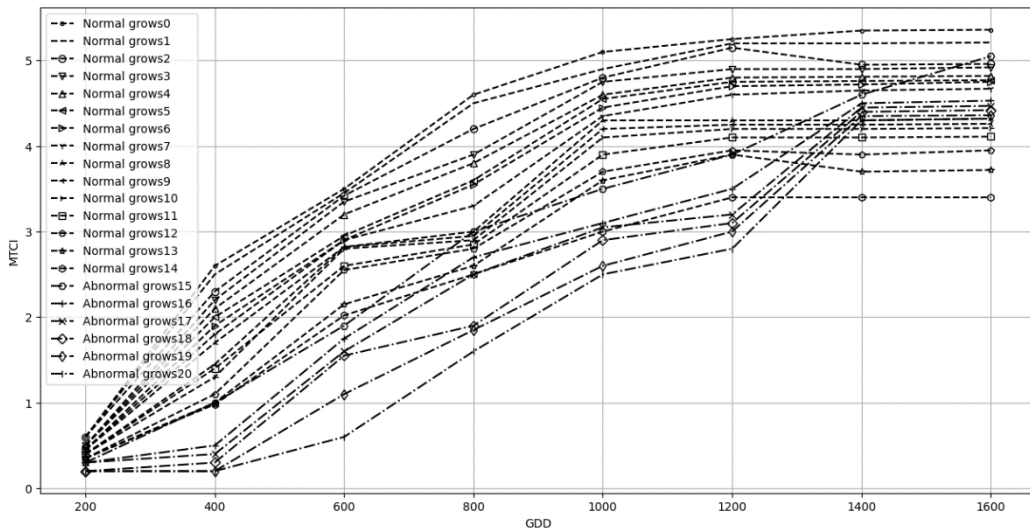


Рис. 11. Часові профілі МТСІ

На базі цих даних ми можемо побудувати адаптивну дворівневу модель урожайності. Але попередньо нам потрібно дослідити її адекватність. Тому розбиваємо набори спостережених значень OV на навчальні Tr , тестувальні Ts та контрольні Cn підмножини

$$OV = Tr \cup Ts \cup Cn \quad (7)$$

На основі множини Tr будуюмо адаптивні моделі вегетативних індексів NDVI, МТСІ згідно поданих співвідношень (1)-(18). Далі будуюмо лінійні регресійні моделі урожайності

$$Y_L = C_1 X_{NDVI} + C_2 X_{MTCI} + C_0 + C \quad (8)$$

або в двофакторній моделі випадкового лісу

$$Y_{RF} = RF(X_{NDVI}, X_{MTCI}) \quad (9)$$

які ідентифікуємо на основі множини тестових спостережень Ts . Далі контролюємо якість побудованих моделей на основі множини контрольних точок, які не брали участі у побудові та навчанні моделей. За результатами аналізу ефективності моделей на контрольних точках будуюмо рекомендації щодо їх раціонального використання.

Результати моделювання динаміки NDVI показали достатньо високу точність для обох типів умов вирощування, що демонструється на рис. 12.

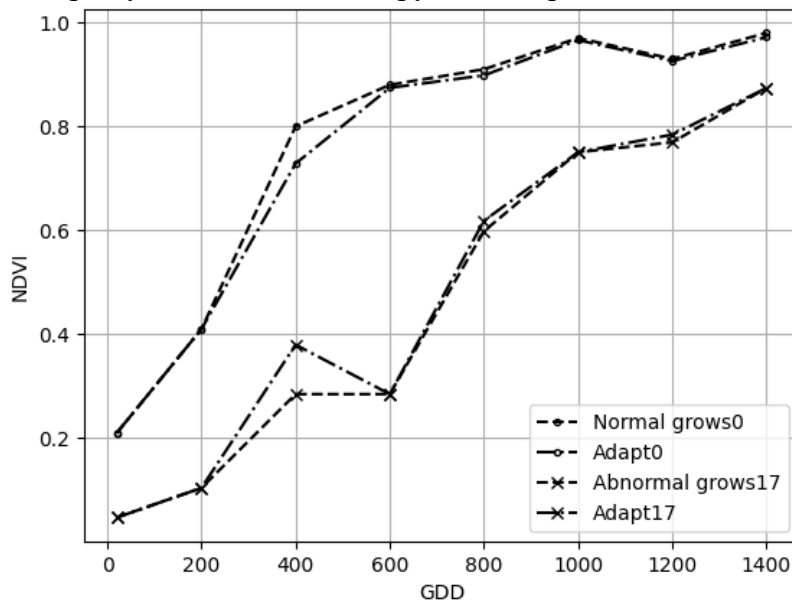


Рис. 12. Адаптивні моделі Моно для NDVI в нормальних та стресових умовах

Моделювання динаміки МТСІ продемонструвало ще вищу точність, особливо для нормальних умов вирощування, як показано на рис. 13.

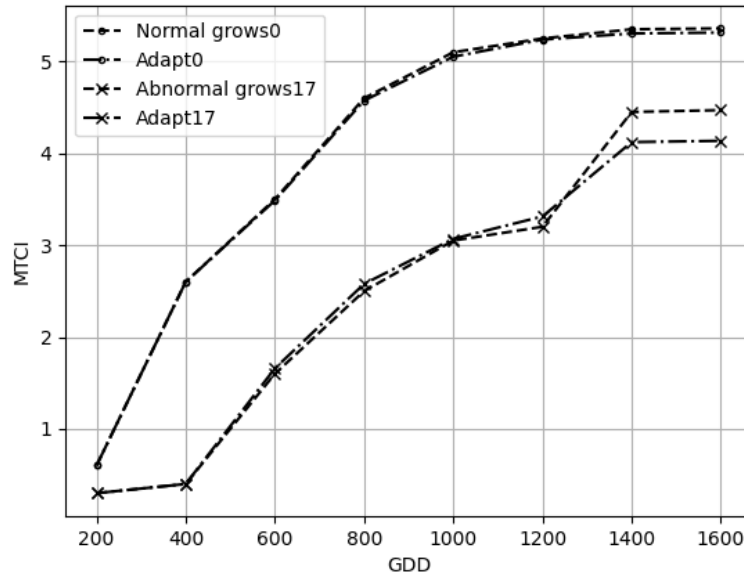


Рис.13. Адаптивні моделі Моно для МТСІ в нормальних та стресових умовах

Узагальнені результати моделювання динаміки вегетаційних індексів демонструють високу ефективність запропонованого двокомпонентного адаптивного підходу. Для NDVI модель показала стабільні результати: у нормальних умовах максимальна похибка 7,3% при середній 1,3%, в стресових умовах – 10,9% та 1,8% відповідно. МТСІ продемонстрував ще вищу точність з мінімальними похибками 0,9%/0,4% у нормальних умовах та 7,5%/2,6% у стресових. Отримані результати підтверджують робастність підходу та його придатність для практичного застосування в системах точного землеробства. Модель лінійної регресії для прогнозування урожайності рису показала помірну точність з чітко вираженою гіперплощиною залежності, що візуалізовано на рис. 14.

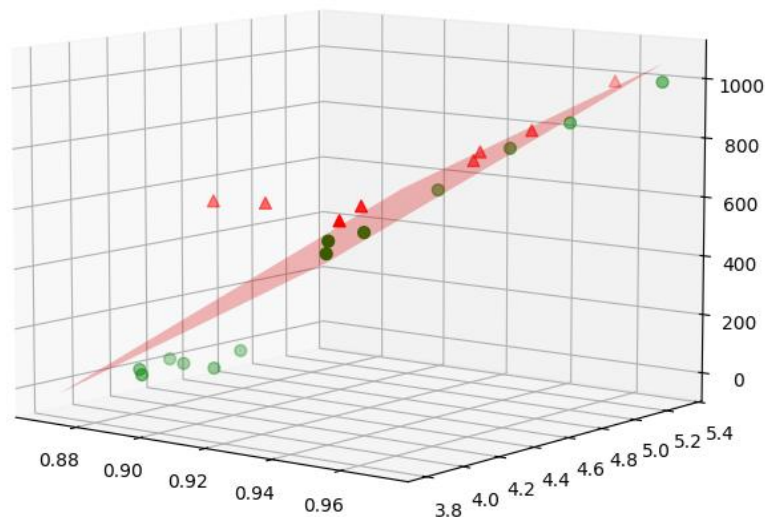


Рис. 14. 3D графік лінійної регресії урожайності

Модель випадкового лісу продемонструвала здатність до кращого відображення нелінійних залежностей між вегетаційними індексами та урожайністю, що показано на рис. 15.

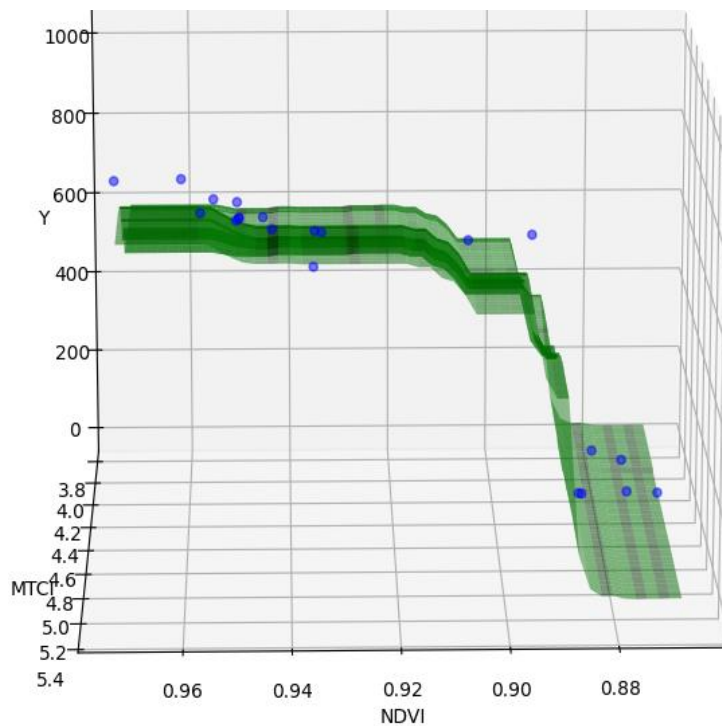


Рис. 15. 3D графік моделі випадкового лісу

Результати двофакторного моделювання урожайності рису підтверджують переваги використання нелінійних методів машинного навчання для прогнозування сільськогосподарської продуктивності. Порівняльний аналіз показав, що модель випадкового лісу демонструє суттєво вищу точність порівняно з лінійною регресією: максимальна похибка знижується з 17,7% до 14,1%, а середня похибка зменшується майже вдвічі з 8,3% до 5,3%. Це свідчить про наявність складних нелінійних взаємозв'язків між вегетаційними індексами NDVI та MTCI та кінцевою урожайністю, які краще відображаються ансамблевими методами. Отримана середня похибка 5,3% для моделі випадкового лісу є цілком прийнятною для практичного застосування в агрономічних системах підтримки прийняття рішень.

Висновки. Вперше запропоновано комплексний підхід до прогнозування урожайності сільськогосподарських культур, який базується на інноваційній двокомпонентній адаптивній моделі динаміки вегетаційних індексів. Ключовим науковим досягненням є успішне поєднання дискретних адаптивних методів з неперервним моделюванням на основі системи диференціальних рівнянь Моно, що забезпечило високу точність відтворення біологічних процесів росту рослин. Розроблена математична модель виявлення ущільнень ґрунту демонструє суттєвий прорив у підвищенні точності діагностики проблемних ділянок поля, забезпечуючи покращення результатів у півтора рази порівняно з традиційними лінійними методами. Це досягнення має важливе практичне значення для оптимізації агротехнічних заходів та підвищення ефективності використання сільськогосподарської техніки. Методологічний внесок роботи полягає у створенні дворівневої системи моделювання, де вегетаційні індекси NDVI та MTCI виступають не як прості предиктори, а як модельовані змінні з власною динамікою. Такий підхід забезпечив високу точність прогнозування з середньою похибкою 5,3% для ансамблевих методів, що є цілком прийнятним для практичного використання в агрономічних системах підтримки прийняття рішень. Технологічною основою реалізації розробленого підходу стала архітектура геоінформаційної системи на базі PostGIS, яка забезпечує ефективну обробку великих масивів просторово-часових даних від БПЛА та LiDAR систем. Експериментальна валідація на реальних агрономічних даних пшениці та рису в різних кліматичних умовах підтверджує універсальність та практичну застосовність запропонованих методів для розвитку систем точного землеробства.

Список літератури

1. Basso B., Liu L. Seasonal crop yield forecast: Methods, applications, and accuracies. *Advances in Agronomy*. 2019. Т.154. С. 201-255.
2. Xue J., Su B. Significant remote sensing vegetation indices: A review of developments and applications. *Journal of Sensors*. 2017. 1353691.
3. Lobell D.B., Burke M.B. On the use of statistical models to predict crop yield responses to climate change. *Agricultural and Forest Meteorology*. 2010. Т.150 (11). С.1443-1452.
4. Reynolds M., Dreccer F., Trethowan R. Drought-adaptive traits derived from wheat wild relatives and landraces. *Journal of Experimental Botany*. 2007. Т.58(2). С. 177-186.
5. Pasichnyk R. M., Babala, L. V., & Machuliak, M. V. A Method for Improving the Quality of Image Annotation in Semantic Monitoring Gis of Business Processes. *Informatics and Mathematical Methods in Simulation*. 2024. Т.14(3).
6. Pasichnyk R., Babala L., Machulyak M. Vegetation Indices Dynamics Model in GIS Based on an Adaptive Predictive Method and the Mono System. *15th International Conference on Advanced Computer Information Technologies (IEEE)* . 2025. P. 186-191).
7. Jones J.W. The DSSAT cropping system model. *European Journal of Agronomy*. 2003. V.18 (3-4). P. 235-265.
8. Zhang W., Liu L., Wang T. Deep learning approaches for crop yield prediction using satellite imagery. *Remote Sensing*. 2024. V.16 (4). P. 875.
9. Monod J. Recherches sur la croissance des cultures bactériennes. Paris: Hermann & Cie, 1942

**TWO-COMPONENT ADAPTIVE MODEL OF VEGETATION INDICES DYNAMICS
FOR AGRICULTURAL CROP YIELD PREDICTION**

M.V. Machulyak

Western Ukrainian National University
11, Lvivska Str., Ternopil, 46009, Ukraine
Email: Mvmach9@gmail.com

The paper proposes a two-component adaptive model of vegetation indices dynamics, which, unlike existing ones, contains a discrete adaptive component based on a combination of current and historical data and a continuous approximation component based on the Monod differential equations system. A method for identifying a two-level adaptive yield model has been developed that uses explanatory variables in the form of vegetation indices models NDVI and MTCI instead of their direct values. The proposed geographic information system architecture based on PostGIS provides efficient integration of spatial data from UAV and LiDAR systems. A mathematical model for detecting soil compaction areas based on Monod equations showed a 1.5-fold accuracy improvement compared to linear methods. Experimental validation on wheat and rice data confirmed the approach effectiveness with an average yield prediction error of 5.3% for ensemble methods. The obtained results demonstrate the promising application of the proposed approach for precision agriculture systems development and agricultural decision-making support.

Keywords: vegetation indices, Monod model, yield prediction, adaptive modeling, remote sensing, geographic information systems, precision agriculture