

ВИКОРИСТАННЯ АЛГОРИТМУ ЛІНІЙНОЇ РЕГРЕСІЇ ДЛЯ ПРОГНОЗУВАННЯ ВРОЖАЙНОСТІ ПШЕНИЦІ СОРТУ «ОКТАВА ОДЕСЬКА» В УМОВАХ СТЕПОВОЇ ЗОНИ МИКОЛАЇВСЬКОЇ ОБЛАСТІ

ПРИСТАШ С.Ф. – кандидат технічних наук

orcid.org/0009-0001-6271-7395

Миколаївський національний аграрний університет

ПРИСТАШ М.С. – кандидат технічних наук

orcid.org/0000-0002-7617-6200

Інститут імпульсних процесів і технологій Національної академії аграрних наук України

Постановка проблеми. Традиційні методи прогнозування врожайності сільськогосподарських культур є трудомісткими та часто неточними. Фермери та агрономи використовують різноманітні фактори, такі як погодні умови, тип ґрунту та сорти рослин, але точний прогноз врожаю залишається складним завданням. Сучасні технології моделювання пропонують більш точні та ефективні способи прогнозування врожаїв, дозволяючи фермерам оптимізувати використання ресурсів та підвищити ефективність виробництва.

Сучасні методи моделювання врожайності можна розділити на дві основні групи:

1. Моделювання на основі фізичних моделей: Цей підхід передбачає створення математичних моделей, які описують взаємодію рослин з навколишнім середовищем. Фактори, такі як кількість опадів, температура, рівень сонячного світла та склад ґрунту, враховуються в цих моделях для прогнозування росту рослин та врожайності.

Переваги: Точні результати за умови правильного розуміння фізичних процесів, можливість адаптації до різних культур та умов.

Недоліки: Складність створення моделей, необхідність великої кількості даних, обмежена здатність враховувати всі фактори, що впливають на врожайність.

2. Моделювання на основі даних (машинне навчання): Цей підхід використовує великі обсяги історичних даних про погоду, ґрунт, добрива, шкідників та інші фактори для прогнозування врожайності. Алгоритми машинного навчання аналізують ці дані та виявляють закономірності, які пов'язують різні фактори з врожайністю.

Переваги: Здатність враховувати велику кількість даних та складні взаємозв'язки, висока точність прогнозів за наявності якісних даних.

Недоліки: Моделі можуть бути «чорним ящиком», тобто не завжди зрозуміло, як саме вони приймають рішення, залежність від якості даних, можливість пере-навчання.

Обидва підходи мають свої переваги та недоліки. Оптимальний вибір методу залежить від конкретних умов та поставлених завдань. Комбінування фізичних моделей та машинного навчання може забезпечити більш точні та надійні прогнози врожайності.

Використання технологій моделювання в аграрному секторі дозволяє:

- Оптимізувати використання ресурсів: добрив, води, пестицидів.

- Зменшити ризики, пов'язані з несприятливими погодними умовами та шкідниками.

- Підвищити ефективність виробництва та прибутковість сільськогосподарських підприємств.

Таким чином, моделювання врожайності є важливим інструментом для сучасного сільського господарства. Воно дозволяє приймати більш обґрунтовані рішення та підвищувати продуктивність сільськогосподарських культур.

Аналіз останніх досліджень. Моделювання врожайності м'якої озимої пшениці сорту «Октава одеська» [1-2] в степах Миколаївської області є актуальним завданням сучасного агровиробництва. Зміна кліматичних умов, виснаження ґрунтів та поява нових шкідників вимагають постійного вдосконалення технологій вирощування.

Традиційні методи прогнозування врожайності часто не враховують великої кількості факторів, що впливають на рівень урожаю. Тому використання методів машинного навчання, які дозволяють аналізувати великі обсяги даних та виявляти складні залежності, відкриває нові можливості для підвищення точності прогнозів.

Метою цієї роботи є дослідження можливості використання моделей машинного навчання, які б дозволили прогнозувати врожайність пшениці сорту «Октава одеська» в умовах степової зони Миколаївської області з урахуванням таких факторів, як температура та опади.

Матеріали та методика досліджень. В загальному вигляді процес машинного навчання можна відобразити схемою, зображеною на Рис. 1. Він складається з вхідних даних, алгоритму, та результатів, які були отримані. Додатково до процесу можуть бути додані данні попереднього «навчання» моделі що дозволяє оцінити точність прогнозу. Така модель має назву «контрольована».

Вхідні данні: $X_1, X_2 \dots X_n$

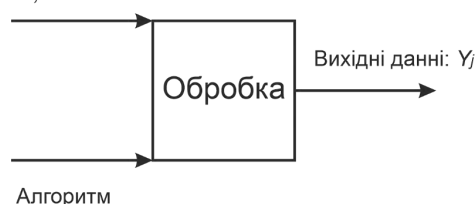


Рис. 1. Загальна схема моделювання за допомогою машинного навчання

В якості алгоритмів модулювання було обрано лінійну регресію. В якості мови реалізації – язык програмування Python та додаток Colaboratory, або скорочено «Colab» – продукт від Google Research. Colab, щоб писати та виконувати довільний код Python через браузер, і який особливо добре підходить для машинного навчання та аналізу даних [3].

Логістична регресія – це алгоритм класифікації, де взаємозв'язок між даними моделюється за допомогою лінійних функцій, а невідомі параметри моделі оцінюються за вхідними даними. Це один з найпростіших алгоритмів, який можна використовувати для вирішення задач класифікації [4-5], що одночасно є як його перевагами, так і недоліками.

В якості даних для моделювання були використані результати врожайності озимої пшениці сорту «Октава одеська», отримані в період з 2020 року до 2024 року та представлені у відкритому доступі в репозитарії Миколаївського національного аграрного університету [2]. Значення температури та кількості опадів в середмісті м. Миколаїв було взято з сайту Українського гідрометеорологічного центру [6] та сайту Метеопост [7].

Для опису отриманих в результаті даних було використано рівняння лінійної регресії та поліноміальної регресії другого ступеня. Поліноміальна регресія – це форма регресійного аналізу, в якому залежність між незалежною змінною x і залежною змінною y моделюється як поліном від x ступеню n . Цей метод використовується для моделювання нелінійних залежностей.

Для лінійної регресії з двома ознаками (X , Y), рівняння буде виглядати так:

$$N = a \cdot X + b \cdot Y + c \quad (1)$$

де a та b – це коефіцієнти для відповідних ознак, а c – це константа.

Для поліноміальної регресії другого ступеня з двома ознаками (X , Y)

$$N = a \cdot X^2 + b \cdot Y^2 + c \cdot X + d \cdot Y + e \cdot X \cdot Y + f \quad (2)$$

де a , b , c , d , e – це коефіцієнти для відповідних термінів, а f – це константа.

Для прогнозування вихідних даних була використана контрольована модель машинного навчання. Для визначення точності прогнозу та при отриманні коефіцієнтів регресії використовувались наступні метрики [8–10]:

– Середня абсолютна помилка (Mean Absolute Error, MAE) – це міра помилок між спареними спостереженнями, що виражають одне й те ж явище. MAE розраховується як сума абсолютних помилок, поділена на розмір вибірки. Це просто середнє абсолютне відхилення між X і Y . Кожна помилка вносить у MAE внесок, пропорційний абсолютному значенню помилки.

– Середньоквадратична помилка (Mean Squared Error, MSE) – це міра середнього з квадратів помилок, тобто середнього квадратичного відхилення між оціненими значеннями та справжнім значенням. MSE завжди додатна, і значення 0 (яке майже ніколи не досягається на практиці) вказувало б на ідеальну відповідність даним. Менший MSE кращий, ніж більший.

Результати досліджень. На рисунку 2 наведено дані щодо середньої температури, опадів та врожайності озимої пшениці сорту "Октава одеська" в період з 2020 року до 2024 року. У роки, коли середня температура взимку (синя лінія) перебувала у певному оптимальному діапазоні (0–15 °C), врожайність (червона лінія) була максимальною. У роки з нижчою або вищою температурою взимку спостерігалися зниження врожайності.

Багато досліджень [11–14] підтверджують, що саме зимовий період є критичним для формування потенціалу

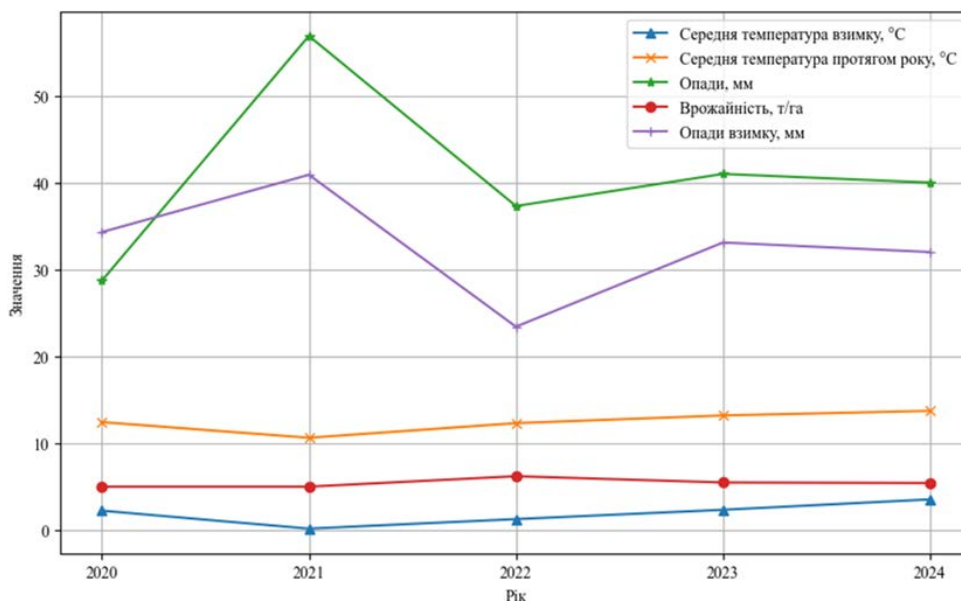


Рис. 2. Показники середньої температури, опадів та врожайності озимої пшениці сорту «Октава одеська» в період з 2020 року до 2024 року

озимої рослини. За низьких температур метаболізм уповільнюється, а при надто високих виникає посуховий стрес і надмірне висихання ґрунту, що ускладнює утворення зв'язі. Річна середня температура впливає опосередковано: за відсутності екстремальних коливань вона може компенсувати незначні відхилення від оптимуму, але ключовими залишаються саме зимові умови.

Достатня й рівномірно розподілена протягом року волога підвищує врожайність, тоді як надлишок або дефіцит опадів порушує водний баланс рослини. Особливо важливими є опади взимку: оптимальна їх кількість забезпечує потрібний ґрунтовий запас вологи для формування зав'язі, а надлишок чи нестача – погіршує розвиток колосків.

Отже, середньозимова температура та зимові опади мають найпряміший вплив на врожайність сорту «Октава одеська»: оптимальні показники цих факторів створюють сприятливі умови для закладки й початкового росту колосків.

У побудованій лінійній регресійній моделі (Рис. 2) врожайність сорту «Октава Одеська» залежить від двох кліматичних змінних – середньозимової температури і кількості опадів у зимовий період. Тривимірна поверхня прогнозованих значень – плоска, що свідчить про адитивну й лінійну природу моделі: підвищення середньозимової температури на кожен градус дає приріст врожайності приблизно на 0,08 т/га, тоді як кожен додатковий міліметр опадів узимку спричиняє її зниження на $\approx 0,06$ т/га. Найвищі прогнозовані врожаї (близько 5,7 т/га) досягаються за поєднанням теплих зим із температурами $+8...+10^{\circ}\text{C}$ та помірно низьких опадів (10–20 мм), водночас найнижчі (близько 0,5–1 т/га) – у холодні ($-10...-8^{\circ}\text{C}$) і вологі

(70–90 мм) зими. Значення середньоквадратичної похибки ($\text{MSE} = 0,2176$) та середньої абсолютної похибки ($\text{MAE} = 0,4665$) свідчать, що прогнозовані дані лежать доволі близько до фактичних спостережень, з середньою помилкою менше півтони на гектар. Отримані результати підтверджують прямий позитивний вплив м'яких зим на формування потенціалу культури та негативний ефект надмірного зволоження ґрунту у холодний період, що дозволяє практично використовувати модель для оцінки врожайності за прогнозованих кліматичних умов та планування агротехнічних заходів.

В результаті апроксимації поверхні залежності врожайності від опадів та температури взимку за допомогою лінійного рівняння (Рис. 4, а), було отримано схожу поверхню, та коефіцієнти а, b, та с, що дорівнюють 0,08, -0,06 та 6,78 відповідно. Значення середньоквадратичної похибки ($\text{MSE} = 0,06$) та середньої абсолютної похибки ($\text{MAE} = 0,16$), що свідчить про високу точність моделі.

Завдяки нелінійним термінам поверхня апроксимації за допомогою поліному другого ступеню (Рис. 4, б) має легку увігнутість: зростання температури спричинює поступове збільшення врожайності з ефектом насичення при вищих значеннях, тоді як вплив опадів узимку спочатку суттєво знижує врожайність, але згодом його негативний ефект вирівнюється.

Найвищі мітки (жовті точки) зосереджені там, де температура вища ($+8...+10^{\circ}\text{C}$) при помірних опадах (20–40 мм), а найнижчі (фіолетові) – за холодних ($-10...-8^{\circ}\text{C}$) і вологих (70–90 мм) зим. На відміну від лінійної моделі, поліномна апроксимація краще відтворює ці нелінійні взаємодії: MAE знизилася до 0,1136, а MSE – до 0,0227, що свідчить про високу точність

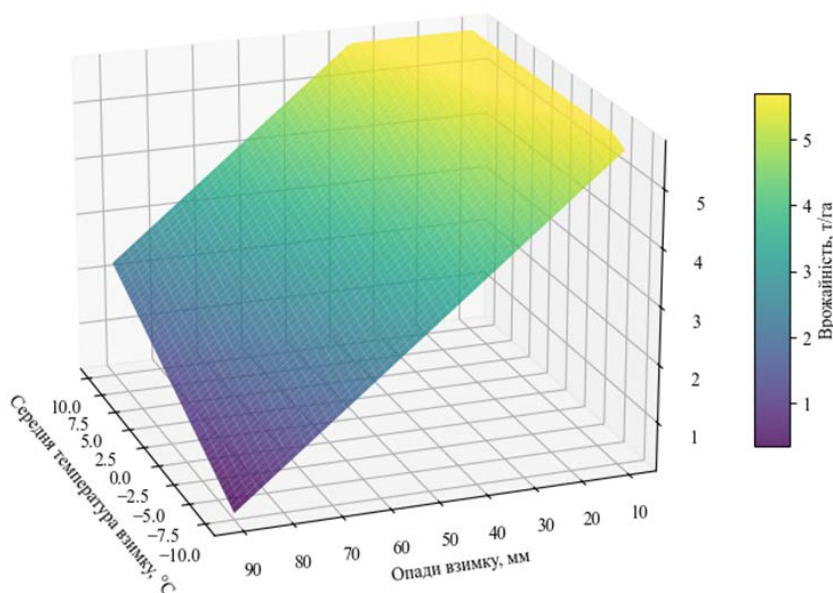
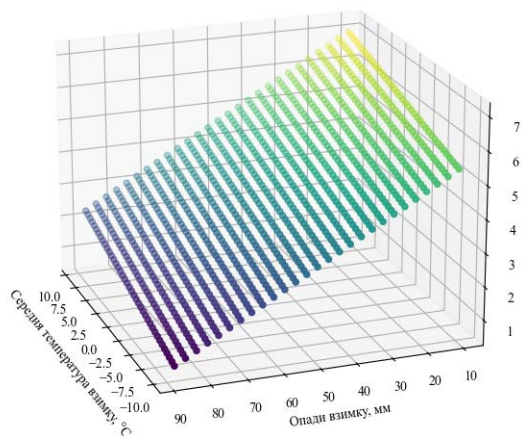
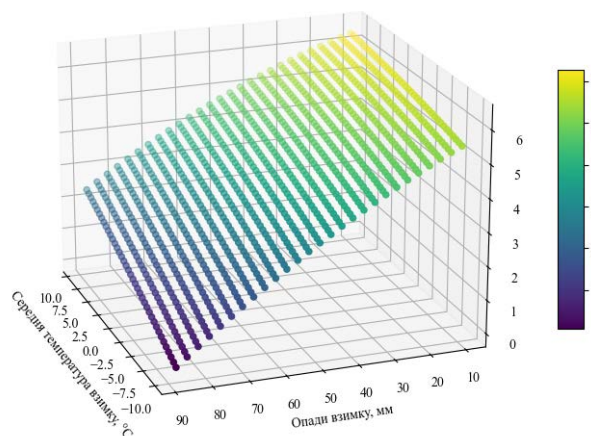


Рис. 3. Розподіл значень врожайності в залежності від середньої температури та опадів взимку отриманий за допомогою алгоритму лінійної регресії. Максимально можливе значення врожайності не перевищує 5.7



а



б

Рис. 4. Результати апроксимації, отриманих за допомогою алгоритму лінійної регресії результатів, за допомогою лінійного рівняння (а) та за допомогою поліному другого ступеню (б)

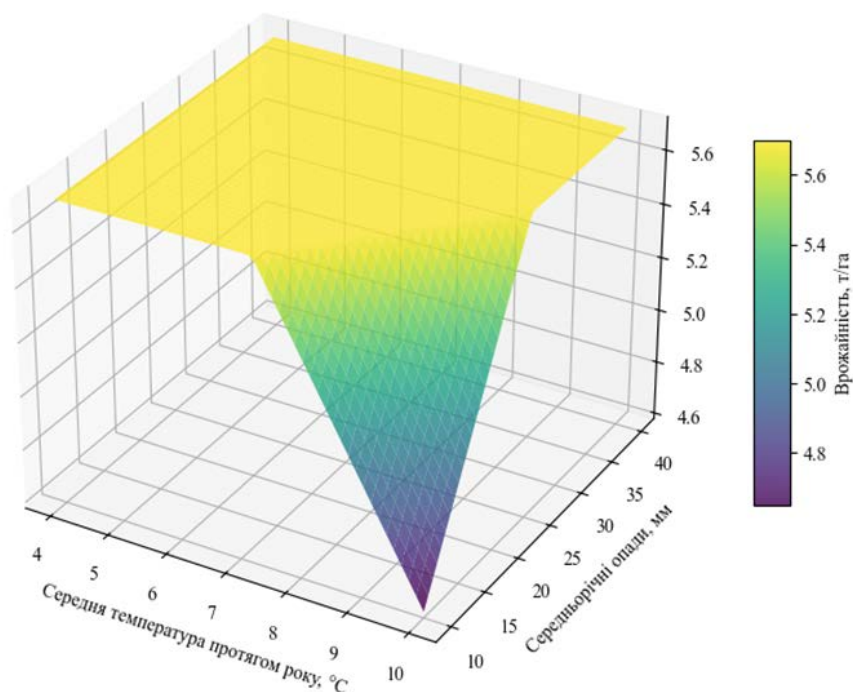


Рис 5. Розподіл значень врожайності в залежності від середньої температури та опадів протягом року отриманий за допомогою алгоритму лінійної регресії. Максимально можливе значення врожайності не перевищує 5.7

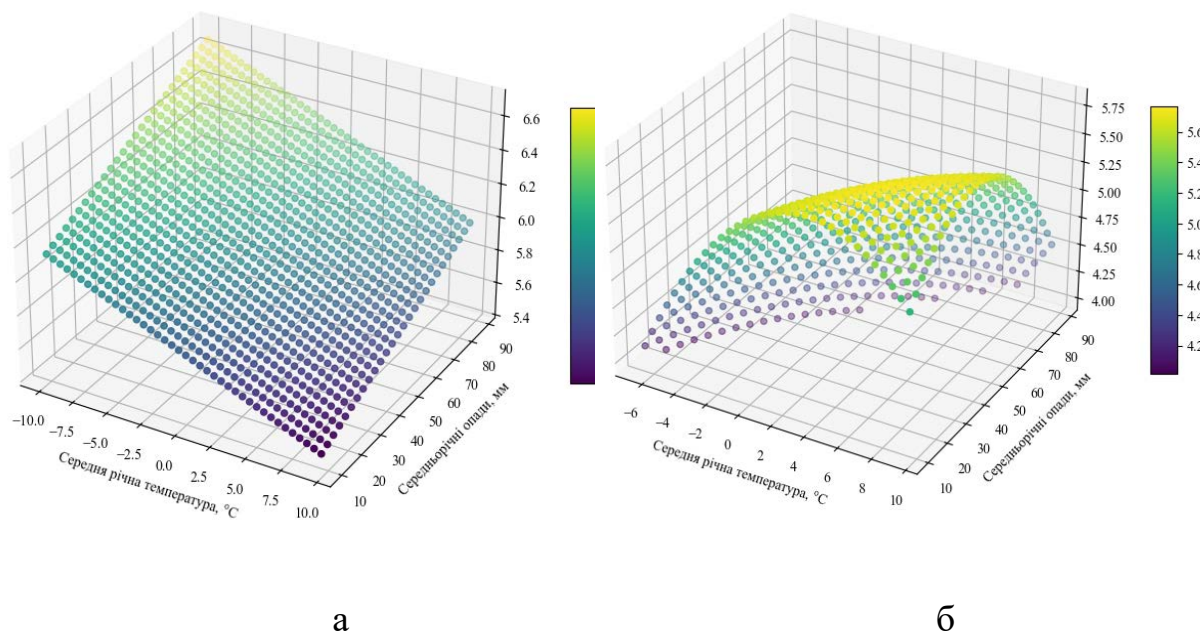


Рис. 6. Результати апроксимації, отриманих за допомогою алгоритму лінійної регресії результатів, за допомогою лінійного рівняння (а) та за допомогою поліному другого ступеню (б)

і менші систематичні відхилення. Коефіцієнти рівняння апроксимації a , b , c , d , e складають: 0,0345, -0,0345, -0,00085, 0,00094, -0,00025, а $f = 6,315$.

На рисунку 5 наведено результати розподілу значень врожайності в залежності від середньої температури та опадів протягом року отримані за допомогою алгоритму лінійної регресії. Чітко простежуються два головні тренди: із ростом середньорічних опадів врожайність підвищується, а зі зростанням температури – падає. У лінійній регресії це реалізовано як плоска поверхня з негативним нахилом уздовж осі температури та позитивним – уздовж осі опадів, обрізана максимумом генетичного потенціалу сорту.

Низька волога (10–15 мм на місяць) призводить до жорсткого водного стресу на всіх фенофазах: зменшується листкова поверхня, сповільнюється фотосинтез, слабшає розвиток кореня – тому прогнозована врожайність опускається до 4,6–5,0 т/га. Кожні додаткові 5–10 мм опадів дають $\approx 0,1$ – $0,2$ т/га надбавки, доки волога не стане надмірною чи не почнуть діяти інші обмежувальні фактори.

По-друге, підвищені температури (понад 8 °C середньорічно) прискорюють вегетацію, зменшують тривалість закладки колосу та зернівки, збільшують випаровування та ризик теплового стресу в критичні фази (колосіння–наліт). Тому при зміні температури від 4 до 10 °C модель прогнозує втрату $\approx 1,1$ т/га ($\approx 0,18$ т/га на кожний 1 °C), що корелює з укороченням періоду наливу зерна.

Нарешті, «плато» максимальних значень (жовта ділянка) відображає зону, де ні волога, ні тепло вже не є обмежувачами факторами – далі вихід на граничний потенціал сорту стримується генетикою.

При апроксимації отриманої поверхні лінійною регресією (Рис. 6, а), результат відображає просторову

площину з негативним нахилом за температурною віссю та позитивним – за опадовою, без «зрізу» максимального потенціалу. Так, кожен додатковий градус у середньорічній температурі зменшує врожайність на приблизно 0,02–0,03 т/га, а кожні 10 мм опадів додають близько 0,05 т/га. Відсутність випуклих «плато» чи «куполів» підкреслює лінійну природу моделі: чим вища волога й чим нижча температура, тим кращий прогноз. Показники точності (MAE = 0,08 т/га; MSE = 0,018) свідчать про хорошу узгодженість моделі з емпіричними даними, але одночасно попереджають про можливу недооцінку нелінійних ефектів екстремальних умов. Коефіцієнти a , b , та c дорівнюють -0,35, 0,007 та 5,73 відповідно.

На відміну від плоскої лінійної поверхні, поліном другого ступеню формує виразний опукло-вгнутий «купол»: початкове зростання врожайності з підвищенням опадів сповільнюється біля 50–60 мм, після чого приріст стає мінімальним, а надмірна вологість у поєднанні з екстремальними температурами злегка знижує продуктивність. Аналогічно, оптимальний температурний діапазон (приблизно +2...+7 °C) у поєднанні з помірним зволоженням дає найбільші одержані значення врожайності – близько 5,6–5,7 т/га, тоді як відхилення в будь-який бік призводить до поступового зниження прогнозованих показників. Надзвичайно низькі температури (<-2 °C) та низькі опади (<20 мм) формують «відроги» моделі з врожайністю до 4,1–4,3 т/га. Сукупність цих спостережень підтверджується високою точністю апроксимації: середня абсолютна похибка становить MAE = 0,0637; MSE = 0,0078, що свідчить про відмінну відповідність моделі емпіричним даним і про доцільність використання поліноміальної регресії для передбачення врожайності за агрокліматичними факторами.

Коефіцієнти рівняння апроксимації a , b , c , d , e складають: 0,06, 0,0002, -0,016, 0,0049, -0,0006, а $f = 6,56$.

Висновки.

1. Лінійні моделі довели свою ефективність для швидкого прогнозу врожайності («Октава Одеська») з прийнятною точністю ($MAE \leq 0,47$ т/га, $MSE \leq 0,22$), хоча поліноміальні підходи краще відтворюють нелінійні ефекти екстремальних умов. Вони забезпечують оперативну оцінку очікуваного врожаю і можуть інтегруватися в системи раннього попередження агрономів.

2. Отримані рівняння регресій дозволяють кількісно оцінювати вплив температури й опадів, що дає змогу моделювати різні кліматичні сценарії та планувати агротехнічні заходи. Завдяки цьому аграрії можуть розрахувати економічну доцільність додаткового зрошення чи захисту від стресу за конкретних погодних умов.

3. Для максимізації врожайності в Миколаївській області слід підтримувати помірну зимову вологість ($\approx 20\text{--}40$ мм) та температуру ($+2\text{...}+7$ °C), використовувати зрошення при дефіциті опадів і застосовувати поліноміальні моделі за нестандартних кліматичних умов. Регулярний моніторинг місцевих метеоданих і адаптація стратегії під реальні спостереження допоможуть наблизити врожайність до генетичного потенціалу сорту.

СПИСОК ВИКОРИСТАНОЇ ЛІТЕРАТУРИ:

1. Селекційно-генетичний інститут. Національний центр насіннєзнавства та сортівивчення : офіц. веб-сайт / Селекційно-генетичний інститут. – Режим доступу: <https://sgi.in.ua/> (дата звернення: 12.12.2024).
2. Інституційний репозиторій Миколаївського національного аграрного університету : офіц. веб-сайт / Миколаївський нац. аграр. ун-т. – Режим доступу: <https://dspace.mnau.edu.ua/jspui/handle/123456789/9647?mode=full> (дата звернення: 12.12.2024).
3. Prystash M. S., Prystash S. F., Torpakov A. S., Lypian Ye. V., Syzonenko O. M. The use of machine learning methods to predict the processes and results of high-voltage electric discharge processing of titanium powder in kerosene with the implementation of volume-distributed multi-spark discharge. *IOP Conf. Ser.: Mater. Sci. Eng.* – 2023. – Vol. 1277, Art. 012001 – DOI 10.1088/1757-899X/1277/1/012001.
4. Карташов М.В. Імовірність, процеси, статистика : монографія / М.В. Карташов. – Київ : ВПЦ «Київський університет», 2007. – 504 с.
5. Raschka S., Mirjalili V. Python machine learning : machine learning and deep learning with Python, scikit-learn, and TensorFlow 2. *Manchester : Packt Publishing Ltd*, 2019. – 498 p.
6. Український гідрометеорологічний центр Державної служби України з надзвичайних ситуацій : офіц. веб-сайт / Укр. гідрометеоцентр ДСНС. – Режим доступу: <https://www.meteo.gov.ua/> (дата звернення: 12.12.2024).
7. Метеопост. Прогноз погоди в Україні : веб-сайт / Метеопост. – Режим доступу: <https://meteopost.com/> (дата звернення: 12.12.2024).
8. Powers D. M. W. Evaluation: from precision, recall and f-measure to roc, informedness, markedness & correlation. *Journal of Machine Learning Technologies*. – 2011. – Vol. 2, № 1. – P. 37–63. – DOI 10.9735/2229-3981.
9. Willmott C. J., Matsuura K. Advantages of the mean absolute error (MAE) over the root mean square error (RMSE) in assessing average model performance. *Climate Research*. – 2005. – Vol. 30. – P. 79–82. – DOI 10.3354/cr030079.
10. Mean Squared Error (MSE) : електрон. ресурс / ProbabilityCourse.com. – Режим доступу: <https://www.probabilitycourse.com/> (дата звернення: 12.09.2020).
11. Солодушко М. М. Тривалість осінньої вегетації та врожайність пшениці озимої. *Бюлетень Ін-ту зерн. госп-ва УААН*. – 2011. – № 40. – С. 32–35.
12. Наукові основи агропромислового виробництва в зоні Степу України / ред. кол. : М. В. Зубець (голова) та ін. – Київ : Аграр. наука, 2010. – 986 с.
13. Конопльова Є. Л. Особливості росту та розвитку рослин пшениці озимої у період весняно-літньої вегетації в північному Степу України. *Бюлетень Ін-ту зерн. госп-ва УААН*. – 2013. – № 4. – С. 116–120.
14. Гасанова І. І., Ноздріна Н. Л. Ріст та розвиток рослин пшениці озимої в період весняно-літньої вегетації в Північному Степу. *Вісник аграр. науки Причорномор'я*. – 2014. – № 2 (78). – С. 126–131.

REFERENCES:

1. Seleksiino-henetychnyi instytut. Natsionalnyi tsentr nasinnieznavstva ta sortovyvchennia [National Centre of Seed Studies and Variety Testing] : official website: <https://sgi.in.ua/>. Accessed December 12, 2024. [in Ukrainian]
2. Instytutsiyni repozytorii Mykolaivskoho natsionalnoho ahramoho universytetu [Institutional Repository of Mykolaiv National Agrarian University] : official website: <https://dspace.mnau.edu.ua/>. Accessed December 12, 2024. [in Ukrainian]
3. Prystash M. S., Prystash S. F., Torpakov A. S., Lypian Ye. V., Syzonenko O. M. (2023) The use of machine learning methods to predict the processes and results of high-voltage electric discharge processing of titanium powder in kerosene with the implementation of volume-distributed multi-spark discharge. *IOP Conf. Ser.: Mater. Sci. Eng.* 1277, 012001. DOI 10.1088/1757-899X/1277/1/012001.
4. Kartashov M. V. (2007). Imovirnist, protsesy, statystyka [Probability, Processes, Statistics]. Kyiv: VPTs «Kyivskiy universytet», 504 [in Ukrainian]
5. Raschka S., Mirjalili V. (2019) Python machine learning : machine learning and deep learning with Python, scikit-learn, and TensorFlow 2. *Manchester : Packt Publishing Ltd*, 498.
6. Ukrainskyi hidrometeorologichnyi tsentr Derzhavnoi sluzhby Ukrainy z nadzvychainykh sytuatsii [Ukrainian Hydrometeorological Centre of the State Emergency Service of Ukraine] : official website: <https://www.meteo.gov.ua/>. Accessed December 12, 2024. [in Ukrainian]
7. Meteopost. Prohnoz pohody v Ukraini [Weather Forecast in Ukraine] : website: <https://meteopost.com/>. Accessed December 12, 2024. [in Ukrainian]
8. Powers D. M. W. (2011) Evaluation: from precision, recall and f-measure to roc, informedness, markedness & correlation. *Journal of Machine Learning Technologies*, 2(1), 37–63. DOI 10.9735/2229-3981.
9. Willmott C. J., Matsuura K. (2005) Advantages of the mean absolute error (MAE) over the root mean square

- error (RMSE) in assessing average model performance. *Climate Research*, 30, 79–82. DOI 10.3354/cr030079.
10. Mean Squared Error (MSE) (2020) [ProbabilityCourse.com.]: website: <https://www.probabilitycourse.com/>. Accessed December 11, 2024.
 11. Solodushko, M. M. (2011). *Tryvalist osinnoi vehetatsii ta vrozhaunist pshenytsi ozymoi* [Duration of Autumn Vegetation and Yield of Winter Wheat]. *Biuleten Instytutu Zernovoho Hospodarstva UAAN*, 40, 32–35. [in Ukrainian]
 12. Zubets, M. V. (editor) et al. (2010). *Naukovi osnovy ahropromysloвого vyrobnytstva v zoni Stepu Ukrainy* [Scientific Foundations of Agro-industrial Production in the Steppe Zone of Ukraine]. Kyiv: Agrar. nauka, 986 [in Ukrainian].
 13. Konopliova, Ye. L. (2013). *Osoblyvosti rostu ta rozvytku roslyn pshenytsi ozymoi u period vesniano-litnoi vegetatsii v pivnichnomu Stepu Ukrainy* [Features of growth and development of winter wheat plants during the spring-summer vegetation period in the northern Steppe of Ukraine]. *Biuleten Instytutu Zernovoho Hospodarstva UAAN*, 4, 116–120 [in Ukrainian].
 14. Hasanova, I. I., Nozdryna, N. L. (2014). *Rist ta rozvytok roslyn pshenytsi ozymoi v period vesniano-litnoi vegetatsii v pivnichnomu Stepu* [Growth and Development of Winter Wheat Plants during Spring-Summer Vegetation in the Northern Steppe]. *Visnyk Aharnoї Nauky Prychornomoria*, 2(78), 126–131 [in Ukrainian].

Присташ С.Ф., Присташ М.С. Використання алгоритму лінійної регресії для прогнозування врожайності пшениці сорту «Октава одеська» в умовах степової зони миколаївської області

У статті досліджується застосування моделей машинного навчання для прогнозування врожайності озимої пшениці сорту «Октава одеська» в умовах степової зони Миколаївської області. Обґрунтовано необхідність заміни традиційних багатфакторних фізичних моделей більш гнучкими алгоритмами аналізу великих даних. Як вихідні дані використовувалися показники врожайності за 2020–2024 рр., середньозимова та середньорічна температура і кількість опадів, отримані з відкритих джерел Укргідрометцентру та Миколаївського національного аграрного університету. Моделі лінійної регресії й поліномної регресії другого ступеня реалізовані в середовищі Google Colaboratory на Python, що забезпечило швидке прототипування й верифікацію результатів за метриками MAE та MSE.

Лінійна регресія продемонструвала стабільні адитивні зв'язки: з кожним підвищенням середньозимої температури врожайність збільшується на $\approx 0,08$ т/га, а з кожним міліметром зимових опадів знижується на $\approx 0,06$ т/га (MAE = 0,47; MSE = 0,22). Поверхня прогнозів є плоскою з негативним нахилом за температурною віссю і позитивним – за опадною, обрізаною максимумом генетичного потенціалу сорту ($\sim 5,7$ т/га).

Поліномна модель другого ступеня виявила виразні нелінійні ефекти насичення й «плато»: врожайність зростає із збільшенням вологості до 50–60 мм та

температурою до $+2 \dots +7$ °C із подальшим уповільненням приросту за екстремальних значень (MAE = 0,0637; MSE = 0,0078). Такий підхід краще відтворює взаємодію кліматичних факторів у критичних фенофазах.

Отримані залежності дають змогу виконувати сценарне прогнозування врожаю для різних кліматичних умов, оперативно планувати агротехнічні заходи – зрошення, захист від теплового стресу, коригування термінів посіву – та наближати фактичні показники до генетичного потенціалу сорту «Октава одеська». Додатково проведено аналіз чутливості моделі до відхилення вхідних параметрів, що підтверджує її стійкість і практичну застосовність у змінних кліматичних умовах.

Ключові слова: моделювання врожайності, сільське господарство, машинне навчання, фізичні моделі, прогнозування, оптимізація.

Prystash S.F., Prystash M.S. Use of a linear regression algorithm to predict the yield of “Oktava odeska” wheat in the steppe zone of mykolaiv region

The article examines the application of machine learning models to forecast the yield of winter wheat variety “Oktava Odeska” under the steppe conditions of the Mykolaiv region. It substantiates the need to replace traditional multifactor physical models with more flexible big data analysis algorithms. As input data, yield figures from 2020–2024, average winter and annual temperatures, and precipitation amounts were used, obtained from the open sources of the Ukrainian Hydrometeorological Center and Mykolaiv National Agrarian University. The linear regression and second-degree polynomial regression models were implemented in the Google Colaboratory Python environment, enabling rapid prototyping and validation of results using MAE and MSE metrics.

Linear regression demonstrated stable additive relationships: with each degree increase in average winter temperature, yield increases by approximately 0.08 t/ha, while each millimeter of winter precipitation reduces yield by about 0.06 t/ha (MAE = 0.47; MSE = 0.22). The forecast surface is flat, with a negative slope along the temperature axis and a positive slope along the precipitation axis, truncated by the genetic potential maximum of the variety (~ 5.7 t/ha).

The second-degree polynomial model revealed pronounced nonlinear saturation and “plateau” effects: yield increases with moisture up to 50–60 mm and temperature up to $+2 \dots +7$ °C, followed by a slowdown in growth under extreme conditions (MAE = 0.0637; MSE = 0.0078). This approach better captures the interaction of climatic factors during critical phenophases.

The derived relationships enable scenario-based yield forecasting for various climatic conditions, prompt planning of agronomic measures—irrigation, heat stress protection, adjustment of sowing dates—and bring actual results closer to the genetic potential of the “Oktava Odeska” variety. Additionally, a sensitivity analysis of the model to input parameter deviations was conducted, confirming its robustness and practical applicability in variable climatic conditions.

Key words: yield modeling, agriculture, machine learning, physical models, forecasting, optimization.