

АНАЛІЗ МЕТОДІВ І ОСОБЛИВОСТЕЙ ЦИФРОВІЗАЦІЇ ДАНИХ ПОЛЬОВИХ ДОСЛІДЖЕНЬ ЯК БАЗИ ДЛЯ УПРАВЛІННЯ РОСЛИННИЦТВОМ

Волоха М., д-р техн. наук, проф.,

e-mail: volmp@i.ua, <https://orcid.org/0000-0002-0112-7324>

Яблонський П., канд. техн. наук, доц.,

e-mail: ypn@ukr.net, <https://orcid.org/0000-0002-1971-5140>

Лазарчук М.,

e-mail: mlazarchuk@ukr.net, <https://orcid.org/0000-0001-6192-6825>

Лазарчук-Воробйова Ю.,

e-mail: jullazarchuk@gmail.com, <https://orcid.org/0000-0002-7866-3299>

Воробйов О.,

e-mail: voroba.ua@gmail.com, <https://orcid.org/0000-0001-5314-1075>

Грубич М.,

e-mail: Mariya.grubich@gmail.com,

Національний технічний університет України «Київський політехнічний інститут імені Ігоря Сікорського»

Анотація

Мета дослідження. Тема управління рослинництвом, яка значною мірою обумовлюється сучасними цифровізаційними процесами, є актуальною і знаходиться в центрі уваги як спеціалістів та експертів у галузі сільського господарства, так і в сфері обчислювальної техніки, адже виробництво продукції аграрного сектора відіграє життєво важливу роль у світовій економіці. Враховуючи, що традиційні методи обробки польових даних не можуть задовольнити постійно зростаючі потреби сільгоспвиробників на новому етапі розвитку землеробства і є серйозною перешкодою для отримання необхідної інформації, метою статті є проведення критичного огляду й аналізу публікацій із оцифрування баз даних польових досліджень задля розробки й ухвалення ефективних управлінських рішень у рослинництві.

Методи дослідження. Дослідження проводилися з використанням загальноприйнятих наукових методів: абстрактно-логічного; аналізу й синтезу; індукції та дедукції; експертних оцінок.

Результати дослідження. Проведений аналіз дав змогу визначити, що рівень розвитку сільськогосподарських підприємств наразі значною мірою залежить від сучасних цифрових технологій, запровадження яких передбачає зміну загальної парадигми управління виробничими процесами і дає змогу товаровиробникам діяти відповідним чином для збільшення обсягів виробництва. Доведено, що поряд з оновленням матеріально-технічної складової пріоритетом виробництва є інтелектуалізація виробничої та управлінської діяльності на базі цифровізації.

Висновок. З метою отримання вигоди від постійно зростаючого обсягу даних, які надходять із численних джерел цифрової трансформації попри те, що переважна більшість господарників і сільгоспвиробників не є експертами в цій галузі і неспроможні повністю зрозуміти основні закономірності створюваних алгоритмів, запропонований у роботі науково-методичний підхід щодо підвищення ефективності машинного навчання з автоматичного розпізнавання сільськогосподарських культур, виявлення хвороб і бур'янів, прогнозування урожайності і якості врожаю, управління водними ресурсами і ґрунтом може бути корисним для сільськогосподарських підприємств багатьох країн світу.

Ключові слова: машинне навчання, точне землеробство, урожайність, ґрунтові умови, водні ресурси.

Вступ. Постійне збільшення населення Землі та зростаючий попит на продукти харчування поставили перед сучасним сільськогосподарським виробництвом ряд складних завдань. Зокрема, американські вчені різних департаментів відомого Техаського аграрно-технічного університету впевнені, що на екосистеми та нерозривно пов'язане з ними сільське господарство в майбутньому неминуче впливатимуть зміни клімату. Тому, на думку A. Thayer et al. (2020), слід прогнозувати результати змін клімату й обґрунтовувати дії щодо зниження соціальні-екологічних наслідків, розуміння яких допоможе підвищити стійкість екосистеми, необхідну для виробництва продуктів харчування, і зберегти високу якість середовища [Thayer et al., 2020]. Вчені із Саудівської Аравії, Фінляндії та Пакистану стверджують про невідворотність виснаження природних ресурсів і про спотворене ціноутворення на руду та метали, яке підвищує енергоефективність, однак збільшує викиди вуглецю в атмосферу [Nassani et al., 2019]. Науковці Дослідницького центру харчування людини (Гранд-Форкс, США) підкреслюють важливість змін раціону харчування людей із міркувань безпеки для здоров'я, особливо протягом останніх десятиліть [Congrad et al., 2018; Venos et al., 2020]. У нещодавньому систематично-хронологічному огляді концепцій і методів сталого розвитку сільського господарства, проведеному колективом авторів із різних країн на чолі з M. Lampridi, стверджується, що засобом вирішення вищезазначених проблем, що чинять шалений тиск на рівень розвитку сільськогосподарського виробництва, є підвищення ефективності агротехнологій при одночасному зниженні навантаження на навколишнє середовище. Саме ці два вирішальні фактори є основою перетворення виробництва сільськогосподарського сектора економіки на точне землеробство. Така модернізація має значний потенціал для забезпечення стійкості сільськогосподарського виробництва за умови максимальної його продуктивності та безпечною довкілля [Lampridi et al., 2019].

Загалом інтелектуальне сільське госпо-

дарство засноване на чотирьох основних принципах, що дають змогу задовольняти зростаючі потреби населення: а) – оптимальне управління природними ресурсами; б) – використання передових технологій; в) – збереження екосистеми; г) – розвиток адекватних послуг [Zessa, 2019]. Істотною ознакою сучасного сільського господарства, безумовно, є використання інформаційно-комунікаційних технологій (ІКТ), які заохочуються політиками у всьому світі. ІКТ можуть, наприклад, включати інформаційні системи управління господарством (фермою), датчики вологості повітря та ґрунту, акселерометри, мережі бездротових датчиків, камери, дрони, недорогі супутники, онлайн-сервіси, автоматизовані керувані транспортні засоби тощо [Sorensen et al., 2019]. Недоліком і серйозною перешкодою для отримання цінної інформації з даних польових досліджень є те, що традиційні методи обробки даних не можуть задовольнити постійно зростаючі потреби в новій ері «розумного» землеробства на базі створеного алгоритмічного та програмного забезпечення [Evstatiev, Gabrovska-Evstatieva, 2020]. Саме тому з'явилося машинне навчання, яке використовує переваги стрімкого експоненціального зростання обчислювальної потужності й інструменти штучного інтелекту й Інтернету речей [Helm et al., 2020].

Одна з глибинних проблем полягає в тому, що переважна більшість фермерів не є експертами в цій галузі і, отже, неспроможні повністю зрозуміти основні закономірності, отримані за допомогою алгоритмів машинного навчання. Тому слід розробити більш зручні для користувача системи. Зокрема, цінними були б прості для розуміння та експлуатації такі пристрої, як, наприклад, інструмент візуалізації зі зручним інтерфейсом для правильного представлення та обробки даних [Cravero, Serpiveda, 2021; Ang & Seng, 2021]. Враховуючи те, що фермери все більше й більше знайомляться зі смартфонами, індійськими [Mayuri, 2018], й китайськими [Yuan, et al., 2022] вченими запропоновані спеціальні програми для смартфонів як можливе вирішення вищезазначеної проблеми.

Аналіз літературних даних і постановка проблеми. В Україні питанням формування цифровізації як інструмента розвитку інформатизації присвячено багато праць міжнародних та українських науковців і експертів. Аналізуючи діяльність у сфері інформації та телекомунікацій, Irtysheva і Kramarenko (2014) довели необхідність створення трирівневої моделі управління цифровою економікою та розрахували ефективність інвестиційних капіталовкладень у цифровізацію цієї галузі. Про необхідність структурної перебудови економіки для переходу до інформаційного суспільства та забезпечення ефективної державної політики у сфері фінансування інноваційних проєктів зазначає Shpak (2017). Azmuk N. до основних факторів формування людського капіталу відносить мотивацію та самореалізацію, знання іноземних мов, віртуальне співробітництво, професійну відповідальність [Azmuk, 2014]. На думку Т. Lazorenko, І. Sholom (2018), діджиталізація вимагає системного підходу до цифрової трансформації формування людського капіталу. Нгуьбіненко О. під діджиталізацією вбачає покращення комунікаційної взаємодії зі споживачами при виконанні бізнес-процесів [Нгуьбіненко, 2018]. Sokolova Н. (2018) досліджувала аспекти розвитку версій аналогових речей і паперових документів, відео- та фотозображень, звуків. Teteryatnyk В. вважає, що діджиталізація є процесом приведення вихідних даних для провадження господарської діяльності елементів до цифрової форми, заснованим на можливостях сучасної ІТ-індустрії [Teteryatnyk, 2017]. Перетворення паперових книжок у електронні, фотографій – у зображення на екрані тощо є процесом перенесення інформації у цифрову форму [Pyshchulina et al., 2018; Ligonenko et al., 2018]. Gurenko А., Gaschutina О. (2018) пов'язують поліпшення системи обслуговування клієнтів із процесом одночасного використання цифрових технологій.

Отже, більшість авторів, метою досліджень яких була сутність поняття діджиталізації, обмежується процесами розвитку національної економіки в умовах глобалі-

зації. На противагу цьому в статті V. Varga (2020) здійснено контент-аналіз основного терміну «діджиталізація», викладене власне бачення поняття діджиталізації як одного з чинників конкурентоспроможності підприємства, коли завдяки зручнішому використанню цифрових технологій прискорюється робота зі значною базою інформації задля зростання економіки. Trushlyakova А. (2018) зазначає про необхідність досконалого планування цифрової стратегії, вибору ефективної моделі з метою отримання очікуваного ефекту від діджиталізації.

Стосовно процесу цифровізації в агропромисловому комплексі України Rudenko М. (2019), розкриваючи цифровізацію як складову методів сучасного управління в сільському господарстві, відмічає відсутність єдиних підходів до тлумачення її сутності. Багатогранність визначення цього поняття в розумінні автора розглядається з різних позицій: а) держави; б) науковців; в) практиків (підприємців); г) суспільства. У статті Voloshchuk (2019) також окреслено проблеми цифровізації агропромислового комплексу та шляхи їхнього вирішення. Volokha М. et al. (2020) прийшли до висновку, що на рівні сільськогосподарських підприємств або фермерських господарств оцифрування отриманих даних дає змогу здійснювати моніторинг виробничих процесів і процесів росту й розвитку рослин; оцінку стану посівів і прогноз урожайності культур; визначення ділянок, що потребують внесення добрив та отрутохімікатів; контроль сівозмін та якість проведення агротехнічних заходів; визначення вологості ґрунту, оцінку снігового покриву, площ вимерзання тощо.

Процеси оцифрування значних масивів польових даних, тобто збір даних польових досліджень, способи їхньої обробки чи зберігання нерозривно пов'язані з поняттям «великі дані» (Big Data), адже фермери, дослідники й інші спеціалісти сільського господарства наразі накопичують великі обсяги даних. За даними ООН, зростаючий у світі за експонентою великий обсяг даних, створюваних цифровими технологіями, вимагає більших можливостей щодо їхнього

зберігання на додатках електронно-обчислюваних машин (ЕОМ) для редагування, аналізу й інтерпретації [Casella, 2000; The Guardian, 2017]. Водночас варто зазначити, що протягом останніх понад 40 років обсяг зберігання даних збільшився вдвічі. Останнє має значний потенціал і мотивує осіб, які ухвалюють рішення, до підвищення інтелектуального рівня [Sonka, 2016].

Мета і завдання дослідження. Метою дослідження є проведення критичного огляду публікацій із оцифрування баз польових даних задля розробки й ухвалення ефективних управлінських рішень у рослинництві.

Для досягнення мети були поставлені такі завдання:

- систематизувати існуючі методи цифровізації даних польових досліджень на базі вимог 5V: обсяг, різноманітність, швидкість, достовірність, цінність;

- визначити базові засади розробки реалістичних рішень, що поєднують експертні знання спеціалістів у галузі обчислювальної техніки та сільського господарства;

- розробити науково-методичний підхід щодо підвищення ефективності машинного навчання з автоматичного розпізнавання сільськогосподарських культур, виявлення хвороб і бур'янів, прогнозування урожайності та якості врожаю, управління водними ресурсами і ґрунтом.

Матеріали та методи дослідження. Для виявлення відповідних досліджень, що стосуються поставленої мети, опрацьовувалася література, оприлюднена, в основному, за останні п'ять років, застосовувався бібліометричний метод, який дав змогу збирати інформацію про публікації для подальшого узагальнення. Використовувалися наукометричні бази, пошукові системи Scopus, Web of Science, Google Scholar, Science Direct тощо. Для дослідження аспектів управління сільським господарством застосовувалися комбінації ключових слів «управління рослинництвом», «управління водними ресурсами» та «управління ґрунтом» в поєднанні зі словосполученням «машинне навчання». Тексти та посилання в іменних статтях

сканувалися, проводився критичний аналіз публікацій. Відповідними вважалися публікації міжнародних та українських науковців і експертів: журнальні статті, доповіді конференцій, глави монографій, огляди, а також магістерські, кандидатські та докторські дисертації.

Процес огляду літератури базувався на методології синтезу всіх доступних критеріїв і дворівневого аналізу відповідних публікацій, що полегшує систематизацію. Таким чином, здійснено системний огляд літератури понад півтори сотень досліджень методів машинного навчання з управління рослинництвом та ухвалення рішень, оцінки точного рослинництва як складової розумного землеробства за останні п'ять років. Вивчення використаних методологій має особливе значення, оскільки не існує рекомендованих стандартів або норм оцінки стійкості методів у галузі рослинництва. Відомі ж методи найчастіше містять інструменти на основі індикаторів, індексів і багатокритеріальності. Часто спостерігається їхнє комбіноване використання, тому чітке розмежування не завжди можливе.

Релевантність пошуку здійснювалася за TF-IDF-методом [12], відповідно до якого оцінкою важливості слів є статистичний показник, тобто відносна величина $tf-idf$, що є добутком двох співмножників: $tf-idf(t, d, D) = tf(t, d) \times idf(t, D)$, де

$$tf(t, d) = \frac{n_t}{\sum_k n_k}, \quad (n_t - \text{число повторень}$$

слова у даному тексті, $\sum_k n_k$ – всього слів у

тексті); $idf(t, D) = \ln\left(\frac{n_c}{\sum_{j=1}^m n_j}\right)$ – зворотна частота слова або кількість появ цього слова в наборі текстів. Щойно релевантне дослідження було визначено, посилення у статті сканувалися, щоб знайти дослідження, виявлені у ході початкового пошуку.

Показниками рівня значущості рецензованих статей, що зазвичай використовуються в аналізованих дослідженнях для забезпечення загального рівня оцінки і залежать від вибору відповідних метрик

зважування значимості характеристик, були такі: – коефіцієнт кореляції (R); – коефіцієнт детермінації (як правило, квадрат коефіцієнта кореляції R^2); – середня абсолютна помилка (САП); – середня абсолютна помилка у відсотках (САПВ) і середньоквадратична помилка (СКП). Математичні вирази для отримання даних показників відомі з теорії точкової оцінки [Casella, 2000]. Відповідно до цієї теорії, низькі значення показників САП, САПВ та СКП означають невелику помилку і, отже, кращу значущість і продуктивність. На противагу прямування R^2 до 1 демонструє кращу значущість моделі, а також те, що крива регресії ефективно відповідає даним досліджень.

Результати. Проведений збір, систематизація та аналіз публікацій із питань методів та особливостей цифровізації польових даних засвідчив, що визначення великих даних постійно й швидко змінюється у зв'язку з їхніми так званими специфічними вимогами «5V»: 1 – обсяг (Volume), 2 – різноманітність (Variety), 3 – швидкість (Velocity), 4 – достовірність (Validity), 5 – цінність (Value) [Meng et al., 2020]. Якщо ще в 2018 р. Національним інститутом стандартів і технологій США (NIST) відмічалось, що великі дані склалися з наборів, які мають характеристики великого обсягу (1), різноманітності (2), швидкості (3) та мінливості (Variability) (4), тобто 4V [Chang, Grady, 2018], то в 2020–2022 рр., окрім цих характеристик, автори [Sun, Scanlon, 2019; Sibo et al., 2022; Choi et al., 2020], підкреслюючи масштаб і складність великих даних, додали до визначення дескриптори високої розмірності, високої складності та невизначеності (ВН), тобто 5V.

Обґрунтування щодо збільшення кількості одиниць V представлено в огляді Sun, Scanlon (2019), де наголошується про вплив технології великих даних і машинного навчання на різні аспекти управління навколишнім середовищем і водними ресурсами, оскільки прогрес методів дистанційного зондування сприяв поширенню великих даних у багатьох галузях

управління довкіллям і водним режимом, серед яких – прогнозування погоди, запобігання стихійним лихам тощо. Великі дані відкривають нові можливості, хоча вимагають, перш за все, нових форм аналітики інформації за умови правильної інтеграції, з якою машинне навчання допоможе повніше розкрити їхній потенціал. В огляді наведені фактичні додатки великих даних машинного навчання з управління довкіллям і водними ресурсами, де авторами акцентується увага на глибокому навчанні, особливо у зв'язку з тим, що ілюстровані дані з управління довкіллям і водними ресурсами потребують застосування аналітичних засобів аналізу великих даних, починаючи від віртуалізації та граничних обчислень і закінчуючи високопродуктивною обробкою. Відмічено, що одним із найскладніших варіантів використання великих даних і глибокого міждисциплінарного навчання є саме дані з управління довкіллям і водними ресурсами. Обсяг приведених у цьому огляді додатків налічує більше, ніж 1000 розглянутих статей, де доводиться, що: 1) – дослідниками даних із управління довкіллям і водними ресурсами значно зміниться планування, проведення та аналіз експериментів; 2) – найефективніше використання переваг великих даних можливе за умови низьких витрат на їхнє очищення; 3) – завдяки надійнішому та швидшому потоку інформації значна ефективність методів глибокого навчання продемонстрована при вирішенні багатьох проблем покращення зусиль із надання допомоги при завданих стихією лихах. Врешті-решт, найбільш цінними великі дані та штучний інтелект стануть тоді, коли покращать рівень міркування та обговорення. Для великих даних машинного навчання з управління довкіллям і водними ресурсами це означає, що глибоке навчання повинне значно покращити навички розуміння ситуацій і змін, які переживає наша планета з погляду впливу на неї людини та умов довкілля [Sun, Scanlon, 2019].

Китайськими вченими-електротехніками Сіньцзянського університету для

підвищення продуктивності системи моніторингу нерівномірно розподіленої температури в інтелектуальних теплицях на основі ідеї паралельних обчислень розроблено алгоритм об'єднання даних, який оптимізує якість ієрархічної мережі бездротових датчиків і підвищує ефективність обчислень за стратегією злиття даних у реальному часі [Sibo et al., 2022]. Мережа трирівнева: – на нижньому рівні збираються та попередньо обробляються дані про температуру теплиці за допомогою покращеного фільтра Калмана; – на середньому рівні виконується об'єднання попередньо зібраних із нижніх датчиків даних методом паралельного алгоритму об'єднання перетинів; – на верхньому рівні за допомогою створеного алгоритму машини для екстремального навчання об'єднуються дані про температуру із середнього рівня шляхом застосування глобального центру злиття з метою найповнішого відображення температури теплиці. Цей алгоритм, підтверджений результатами моделювання, забезпечує ефективність і точність розробленої стратегії злиття даних у режимі реального часу, що значно підвищило продуктивність системи контролю температури [Sibo et al., 2022].

Підтвердженням з інших сфер діяльності людини може бути приклад проведеного американськими вченими систематичного пошуку літератури в PubMed® за темою використання у медицині, переважно в офтальмології, наукових методів дослідження, який продемонстрував, що сучасні методи штучного інтелекту здатні розробити ефективні діагностичні інструменти для виявлення різноманітних захворювань. Лікарі повинні розуміти суть теорій, покладених в основу штучного інтелекту та його вигоди в галузі медицини [Choi et al., 2020].

Отже, характеризуючи вимоги 5V, виділимо таке.

Обсяг. Показник обсягу належить до розміру великих даних. В останнє десятиліття зростання великих даних співпадає зі швидким розвитком ІКТ та Інтернета, а обсяг їхнього розгортання часто вимірю-

ється терабайтами, петабайтами, інколи ексабайтами (кількість інформації 1 ексабайта дорівнює квінтільйону (10^{18}) байт). Для обробки даних за умови масштабування необхідний перехід на платформу розподілених обчислень. Водночас обсяг набору даних є відносним і може мати різні значення у різних додатках.

Різноманітність стосується методів збору даних. Залежно від джерел великі дані спостережень класифікуються на пасивні чи активні, при цьому перші виходять із похідних продуктів цифрових медіа (наприклад, дані GPS із мобільних телефонів), а другі – із запланованих заходів щодо збору даних (польові спостереження або цільове дистанційне зондування). Основна відмінність полягає в тому, що пасивні дані збираються без активної участі користувача і в результаті не пояснюють причини подій. Залежно від формату дані можуть бути як структурованими (покази датчиків та супутникові дані), так і неструктурованими (електронна пошта, відео, аудіо та текст із соціальних мереж), при цьому основна різниця між ними полягає в тому, що неструктуровані дані не можуть бути легко описані з використанням зумовленої моделі у зв'язку з обмеженістю інформаційних ресурсів у випадку неструктурованості даних. Тобто, різноманітність створює додатковий рівень складності при роботі з великими даними. Для зменшення різноманітності даних зазвичай використовуються два методи: віртуалізація даних та об'єднання даних.

Швидкість. Атрибут швидкості може відноситися не тільки до швидкості генерації даних, але й до необхідної для прийому швидкості аналізу. Швидкість генерації даних є прямим наслідком можливостей підключення, повсюдного використання інтелектуальних пристроїв і мереж моніторингу в реальному часі. Набори високошвидкісних даних, які безперервно генеруються різними джерелами, потребують спеціальної обробки. Потік даних визначається як необмежена послідовність подій, що потребує безперервної

поступової обробки.

Мінливість відноситься до варіацій усіх інших атрибутів великих даних, наприклад, варіацій швидкості потоку даних або змін у їхньому значенні, причому останнє часто має місце в даних, отриманих із краудсорсингу – одного зі способів пошуку вирішення проблем із використанням незначної сторонньої допомоги. Мінливість є серйозною проблемою для обробки даних, оскільки багато алгоритмів машинного навчання працюють на основі принципу стаціонарності.

Достовірність. Складова достовірності належить до потенційних проблем неузгодженості та якості таких даних, як часткова відсутність, аномалії та помилки введення. За високорівневими продуктами даних, що випускаються центрами обробки, полягає значна кількість етапів обробки й злиття. Однак, оскільки центри обробки даних постійно оновлюють алгоритми, управління їхніми версіями стає серйозною проблемою прийому даних із декількох джерел. Отже, при розробці платформ збору та аналізу даних необхідним і важливим є захист достовірності інформації.

Обговорення. Згідно з оглядом літератури за останні п'ять років відносно вирощування сільськогосподарських культур, родючості ґрунту та водного режиму розділи, пов'язані з управлінням водними ресурсами та ґрунтом, є менш вивченими, що в сукупності відповідає 20% від загальної кількості статей (по 10% на кожен категорію), а управління рослинництвом становить більшість статей (до 80% від загальної кількості) і поділяється на чотири категорії, в яких використовуються методи машинного навчання:

- розпізнавання врожаю;
- якість урожаю;
- виявлення захворювань і бур'янів;
- прогноз прибутковості.

Автоматичне розпізнавання сільськогосподарських культур привертає значну увагу в декількох наукових галузях, наприклад, систематика рослин, ботанічні сади та відкриття нових видів. Види

рослин можна розпізнавати та класифікувати за допомогою аналізу різних органів, включаючи листя, стебла, плоди, квіти, коріння та насіння. Використання розпізнавання рослин на основі листя є найбільш поширеним підходом щодо конкретних характеристик листя, наприклад, колір, форма й текстура [Liakos et al., 2018]. З більш широким використанням супутників і літальних апаратів особливої популярності набула класифікація сільгоспкультур за допомогою дистанційного зондування [Virnodkar et al., 2020], тобто розвиток комп'ютерного програмного забезпечення та пристроїв обробки зображень призвели до автоматичного розпізнавання та класифікації сільськогосподарських культур, зокрема за ознакою урожайності.

Якість урожаю дуже важлива для ринку і, як правило, пов'язана з ґрунтовими й кліматичними умовами, методами вирощування та характеристиками сільськогосподарських культур тощо. Високоякісна сільськогосподарська продукція зазвичай продається за вигіднішими цінами, що дає фермерам вищі доходи. Наприклад, якість плодів, твердість м'якоті, вміст розчинних сухих речовин і колір шкірки є одними з найпростіших показників зрілості, використовуваних для збирання врожаю яблук [Parageorgiou et al., 2018]. Терміни збирання суттєво впливають на якісні характеристики зібраної продукції як високоцінних культур (винограду, овочів, дерев'яних культур, трав тощо), так і просапних культур. Таким чином, розробка систем підтримки ухвалення рішень може допомогти фермерам і господарникам в ухваленні відповідних управлінських рішень для підвищення якості продукції. Наприклад, вибірко-вий збір урожаю – це метод управління, який може значно підвищити якість. Окрім того, якість урожаю тісно пов'язана з харчовими відходами, що є додатковою проблемою, з якою доводиться стикатися сучасному сільському господарству: якщо врожай відхиляється від бажаної форми, кольору чи розміру, його можуть визнати

некондиційним. Однак якщо застосовувати технології обробки зображень такої викинутої сировини, розроблені на засадах алгоритмів машинного навчання, то можна забезпечити зменшення кількості харчових відходів для підприємств переробної промисловості.

Хвороби виявлялися досвідченими агрономами шляхом проведення польової розвідки. Однак цей процес займає багато часу та базується виключно на візуальному огляді. Нещодавні технологічні досягнення зробили комерційно доступними сенсорні системи, здатні ідентифікувати хворі рослини до того, як з'являться симптоми. Окрім того, за останні декілька років комп'ютерний зір, особливо з використанням глибокого вивчення, досяг значного прогресу, що дає змогу раніше і якісніше розпізнати симптоми хвороби рослини [Zhang et al., 2021].

Бур'яни зазвичай дуже швидко сходять і розвиваються через багате виробництво насіння і довголіття та поширюються великими частинами поля, конкуруючи з сільськогосподарськими культурами за ресурси, включаючи простір, сонячне світло, поживні речовини та доступність води. Окрім того, бур'яни часто з'являються раніше, ніж сільгоспкультури, і їм не доводиться стикатися з природними ворогами, які негативно впливають на культурні рослини. Щоб запобігти зниженню врожайності, боротьба з бур'яном може відбуватися або механічним обробітком, або застосуванням гербіцидів, або ж поєднанням обох агротехнічних прийомів. Однак механічний обробіток інколи важко здійснити, він є малоефективним, якщо не виконується належним чином, що робить застосування гербіцидів найбільш розповсюдженою практикою. Однак використання великих обсягів гербіцидів є дорого й шкідливо для навколишнього середовища, особливо у випадку рівномірного внесення без урахування різниці в щільності розміщення бур'янів площиною поля. Тривале використання гербіцидів із великою ймовірністю зробить бур'яни стійкішими, що призведе до

більш складної та вартісної боротьби проти забур'янення посівів. В останні роки досягнуто значних успіхів при вирішенні проблеми диференціації бур'янів у посівах сільгоспкультур на основі «розумного» землеробства. Це розрізнення можна досягти за допомогою дистанційного або ближнього зондування датчиками, встановленими на супутниках, повітряних і наземних транспортних засобах, а також як наземних, так і повітряних безпілотниках. Однак перетворення даних, зібраних БПЛА, у «зручну» інформацію є складним завданням, оскільки їхній збір і класифікація вимагають чималих зусиль. Алгоритми машинного навчання у поєднанні з технологіями візуалізації або спектроскопією без візуалізації можуть забезпечити диференціацію та локалізацію певних видів бур'янів у режимі реального часу, забезпечуючи точне застосування гербіцидів і планування найкоротшого шляху пропонування у попередньо визначених ділянках замість оприскування всього поля.

Прогнозування врожайності є однією з найважливіших і найскладніших тем у сучасному сільському господарстві. Точна модель може допомогти, наприклад, власникам ферм ухвалювати обґрунтовані управлінські рішення щодо того, які культури вирощувати, аби урожай відповідав існуючим вимогам ринку. Однак це непросте завдання, воно складається з різних кроків і напрямів цифровізації, адже прогноз урожайності може визначатися декількома факторами, наприклад, технічним рівнем застосовуваних комплексів машин і досконалістю адаптованих технологій, ґрунтово-кліматичними й погодними умовами, генотипічними й фенотипічними характеристиками сільськогосподарських культур та їхньою взаємодією. Отже, це вимагає фундаментального розуміння взаємозв'язку між зазначеними взаємодіючими факторами.

Управління водними ресурсами. Сільськогосподарський сектор є основним споживачем доступної прісної води у глобальному масштабі, оскільки ріст і розвиток рослин значною мірою залежить

від наявності води. Зважаючи на швидке виснаження багатьох водоносних горизонтів використання цифрових технологій забезпечує ефективніше керування процесами накопичення та збереження продуктивної вологи задля забезпечення сталого рослинництва, що також підтверджується дослідженнями вчених Техаського технічного університету [Neurape, Guo, 2019]. У галузі точного землеробства пропонується зрошення зі змінною нормою для економії води [Zhao et al., 2018]. Цього можна досягти, застосовуючи іригаційні норми, які варіюються залежно від мінливості поля на основі конкретних потреб у воді окремих зон управління замість використання єдиної норми для всього поля. Ефективність і використання методу зрошення зі змінною нормою залежать від рівня оцифрування агрономічних факторів, включаючи топографію, властивості ґрунту та їхній вплив на продуктивну вологу ґрунту, щоб досягти як економії води, так і оптимізації врожаю [Lou et al., 2005]. Ретельний моніторинг стану ґрунтової вологи й інших умов росту і розвитку сільгоспкультур у поєднанні з прогнозуванням погодних умов може допомогти у складанні цифрових програм зрошення та ефективному управлінні водними ресурсами, особливо там, де джерела підземних вод використовуються для зрошення, а опади забезпечують лише частину загальної евапотранспірації культур, як досліджено при вирощуванні бавовни в посушливих районах Західного Техасу [Mauget et al., 2017].

Управління ґрунтом. Точна інформація про ґрунт у регіональному масштабі має велике значення, оскільки вона сприяє кращому управлінню ґрунтом відповідно до земельного потенціалу. Ефективність управління ґрунтом також становить значний інтерес у зв'язку з такими проблемами, як деградація земель (втрата біологічної продуктивності), дисбаланс поживних речовин у ґрунті (через надмірне використання добрив) та ерозія ґрунту (внаслідок надмірного зрізання рослинності, неправильної сівозміни, деструктуризації

рушіями енергетичних засобів). Корисні властивості ґрунту можуть включати текстуру, органічну речовину та вміст поживних речовин тощо. Традиційні методи оцінки ґрунту включають взяття проб і лабораторний аналіз, які зазвичай є вартісними та потребують значного часу й зусиль. Водночас датчики дистанційного зондування та картографування ґрунту можуть забезпечити недороге та просте рішення для вивчення його просторової мінливості. Однак злиття й обробка таких різномірних великих даних можуть бути серйозними недоліками при використанні традиційних методів аналізу.

Висновки. Серйозною перешкодою для отримання цінної інформації з польових даних є обмеженість їхньої обробки традиційними методами, особливо на тлі постійно зростаючих потреб точного (розумного) землеробства, що потребує використання інформаційно-комунікаційних технологій на базі алгоритмічного та програмного забезпечення сучасного сільського господарства, яке здатне, наприклад, включати інформаційні системи управління господарством (фермою), датчики вологості повітря та ґрунту, акселерометри, мережі бездротових датчиків, камери, дрони, недорогі супутники, онлайн-сервіси, автоматизовані керувані транспортні засоби тощо. У результаті чого з'явилося машинне навчання, де широко використовуються інструменти штучного інтелекту та Інтернету речей.

Засобом вирішення існуючих проблем, що чинять шалений тиск на рівень розвитку сільського господарства, є підвищення ефективності застосування агротехнологій при одночасному зниженні навантаження на навколишнє середовище на основі перетворення виробництва сільськогосподарського сектора економіки на точне землеробство. Така суттєва і своєчасна модернізація має значний потенціал для забезпечення стійкості сільськогосподарського виробництва за умови його максимальної продуктивності, прибутковості та безпечного довкілля.

На ефективність складних процесів

сільськогосподарського виробництва значний вплив чинить упровадження цифровізації, яка є способом накопичення значних обсягів польових даних та інструментом для розробки нових, сучасних напрямків наукових досліджень.

На основі систематичного аналізу методів і особливостей оцифрування даних польових досліджень задля використання як бази для управління рослинництвом і ухвалення рішень на засадах розробки реалістичних рішень, що поєднують експертні знання спеціалістів у галузі обчислювальної техніки та сільськогосподарства, запропонований науково-методичний підхід щодо підвищення ефективності машинного навчання з автоматичного розпізнавання сільськогосподарських культур, виявлення хвороб і бур'янів, прогнозування урожайності та якості врожаю, управління водними ресурсами й ґрунтом.

References

- Thayer, A., Vargas, A., Castellanos, A., Lafon, C., McCarl, B., Roelke, D., Wine-miller, K., Lacher, T. (2020). Integrating agriculture and ecosystems to find suitable adaptations to climate change. *Climate*, 8, article number 10. <https://doi.org/10.3390/cli8010010>.
- Nassani, A.A., Awan, U., Zaman, K., Hyder, S., Aldakhil, A.M., Abro, M.M.Q. (2019). Management of natural resources and material pricing: Global evidence. *Resources Policy*, 64, article number 101500. <https://doi.org/10.1016/j.resourpol.2019.101500>.
- Conrad, Z., Niles, M.T., Neher, D.A., Roy, E.D., Tichenor, N.E., Jahns, L. (2018). Relationship between food waste, diet quality, and environmental sustainability. *PLoS ONE*, 13(4), article number 0159405. <https://doi.org/10.1371/journal.pone.0195405>.
- Benos, L., Bechar, A., Bochtis, D. (2020). Safety and ergonomics in human-robot interactive agricultural operations. *Bio-systems Engineering*, 200, 55-72. <https://doi.org/10.1016/j.biosystemseng.2020.09.009>.
- Lampridi, M., Sørensen, C., Bochtis, D. (2019). Agricultural sustainability: A review of concepts and methods. *Sustainability*, 11, article number 5120. <https://doi.org/10.3390/su11185120>.
- Zecca, F. (2019). The use of Internet of things for the sustainability of the agricultural sector: the case of climate smart agriculture. *International Journal of Civil Engineering and Technology*, 10, 494-501.
- Sørensen, C. A. G., Kateris, D., Bochtis, D. (2019). ICT innovations and smart farming. *Communications in the field of computer and information sciences*; Springer: Berlin/Heidelberg, Germany, 953, 1-19. https://doi.org/10.1007/978-3-030-12998-9_1.
- Evstatiev, B.I., Gabrovska-Evstatieva, K.G. (2020). Review on the methods for big data analysis in agriculture. *IOP Conf. Series: Materials Science and Engineering*, Borovets, Bulgaria, November 26-29, IOP Publishing Ltd.: Bristol, UK, 1032 p. 012053. <https://doi.org/10.1088/1757-899X/1032/1/012053>.
- Helm, J.M., Svergosh, A.M., Heberley, H.S., Karnuta, J.M., Schaffer, J.L., Krebs, V.E., Spitzer, A.I., Ramkumar, P.M. (2020). Machine learning and artificial intelligence: definitions, applications, and future directions. *Current Reviews in Musculoskeletal Medicine*, 13, 69-76. <https://doi.org/10.1007/s12178-020-09600-8>.
- Cravero, A., Sepúlveda, S. (2021). Use and adaptations of machine learning in big data-applications in real cases in agriculture. *Electronics*, 10, article number 552. <https://doi.org/10.3390/electronics10050552>.
- Ang, K. L. M., & Seng, J. K. P. (2021). Big data and machine learning with hyperspectral information in agriculture. *IEEE Access*, 9, 36699-36718. <https://doi.org/10.1109/access.2021.3051196>.
- Mayuri, K. (2018). The role of image processing methods and machine learning in disease recognition, diagnosis and forecasting of crop yields: A review. *International Journal of Advanced Research in Computer Science*, 9(2), 788-795.
- Yuan, Y., Chen, L., Wu, H., Li, L. (2022). Advanced agricultural disease image recognition technologies: A review. *Infor-*

- mation Processing in Agriculture, 9, 48-59. <https://doi.org/10.1016/j.inpa.2021.01.003>.
- Irtysheva, I., Kramarenko, I. (2014). Investment attractiveness of the economy: interregional asymmetries. *Regional Economy*, 2(72), 84-95.
- Shpak, N. (2017). Features of human capital formation in the agricultural sector of the information economy. *Problems and Prospects of Economics and Management*. 3, 28-38.
- Azmuk, N. (2014). Factors of formation and development of innovative forms of human capital. *Ukraine: Aspects of Work*, 3, 47-51.
- Lazorenko, T., Sholom, I. (2018). Digitalization as the main factor of business development. <http://confmanagement.kpi.ua/proc/article/view/201186>.
- Hrybinenko, O. (2018). Digitalization of the economy in a new paradigm of digital transformation. *International Relations, Part "Economic sciences"*, 16, 35-37.
- Sokolova, H. (2018). Some aspects of the development of the digital economy in Ukraine. *Economic Bulletin of Donbass*. 1(51), 92-96.
- Tetryatnyk, B. (2017). Trends of digitalization and virtualization as a vector of modern development of the world economy. *Scientific Bulletin of the International Humanities University. Series: Jurisprudence*, 29(2), 21-23.
- Pyshchulina, O., Yurochko, T., Mishchenko, M., Zhalilo, Ya. (2018). Development of human capital: on the way to qualitative reforms. *Zapovit, Kyiv, Ukraine*, 367. http://razumkov.org.ua/uploads/article/2018_Lud_Kapital.
- Ligonenko, L., Khripko, A., Domanskii, A. (2018). The content and mechanism of forming a strategy of digitalization in business organizations, *International Scientific Journal "International"*, 22(2), <https://doi.org/10.25313/2520-2057-2018-22-4555>.
- Gurenko, A., Gaschutina, O. (2018). The directions of the management systems development in the conditions of business digitalization of business in Ukraine. *Economy and Society*, 19, 739-745. [32782/2524-0072/2018-19-113](https://doi.org/10.32782/2524-0072/2018-19-113).
- Varga, V. (2020). Digitalization as one of the factors of enterprise competitiveness. *Efektivna ekonomika*, vol. 8. <https://doi.org/10.32702/2307-2105-2020.8.154>.
- Trushlyakova, A. (2018). Development of digitalization in Ukraine: factors of influence, advantages and challenges of today. *Economies Horizons*, 4(7), 186-191. [https://doi.org/10.31499/2616-5236.4\(7\).2018.212762](https://doi.org/10.31499/2616-5236.4(7).2018.212762).
- Rudenko, M. (2019). The features of digitalization influence on the functioning of agricultural enterprises. *Bulletin of the Kharkiv national agricultural university of V. V. Dokuchaeva*. <https://doi.org/10.31359/2312-3427-2019-1-202>.
- Voloshchuk, Y. (2019). Digital directions of agrarian enterprises. *Efektivna ekonomika*, v. 2. <https://doi.org/10.32702/2307-2105-2019.2.66>.
- Volokha, M., Bondar, S., Boldyreva, L., Kisilevskaya, M. (2022). Advantages of digitalization in field data processing and crop management. *Modern research in world science. Proceedings of the 2nd International scientific and practical conference. SPC "Sci-conf.com.ua"*. Lviv, Ukraine. 414-420. URL: <https://sci-conf.com.ua/ii-mezhdunarodnaya-nauchno-prakticheskaya-konferentsiya-modern-research-in-world-science-15-17-maya-2022-goda-lvov-ukraina-arhiv/>.
- Casella, G. (2000). Theory of Point Estimation by E. L. Lehmann. *Journal of the American Statistical Association*, 95(449). <https://doi.org/10.2307/2669560>.
- The Guardian. (2017). Tsunami of Data' Could Consume One Fifth of Global Electricity by 2025. Available online. <https://www.theguardian.com/environment/2017/dec/11/tsunami-of-data-could-consume-fifth-global-electricity-by-2025> (accessed on 11 December 2017).
- Sonka, S. (2016). Big Data: fueling the next evolution of agricultural innovation. *Journal of Innovation Management*, 4(1), 114-136.
- Meng, T., Jing, X., Jan, Z., Pedrycz, W. (2020). Survey on machine learning for data fusion. *Information Fusion*, 57, 115-129.

<https://doi.org/10.1016/j.inffus.2019.12.001>.

Chang, V., Grady. (2018). NIST Big Data Interoperability Framework: Vol. 1, Big Data Definitions, Special Publication. National Institute of Standards and Technology. Gaithersburg, MD. <https://doi.org/10.6028/NIST.SP.1500-1>.

Sun, A.Y., Scanlon, B.R. (2019). How can Big Data and machine learning benefit environment and water management: a survey of methods, applications, and future directions. *Environmental Research Letters*, 14, article number 73001. <https://doi.org/10.1088/1748-9326/ab1b7d>.

Sibo, X., Xinyuan, N., Xin, C., Xumeng, Lu. (2022). Data fusion based wireless temperature monitoring system applied to intelligent greenhouse. *Computers and Electronics in Agriculture*. Vol. 192(C). <https://doi.org/10.1016/j.compag.2021.106576>.

Choi, R.Y., Coyner, A.S., Kalpathy-Cramer, J., Chiang, M.F., Campbell, J.P. (2020). Introduction to machine learning, neural networks, and deep learning. *Translation Vision Science & Technology*, 9, article number 14. <https://doi.org/10.1167/tvst.9.2.14>.

Liakos, K., Busato, P., Moshou D., Pearson, S., Bochtis, D. (2018). Machine Learning in Agriculture: A Review. *Sensors*, 18(8), 2674. doi.org/10.3390/s18082674.

Virnodkar, S., Pachghare, V., Patil, V., Jha, S. (2020). Remote sensing and machine learning for crop water stress determination in various crops: A critical review. *Precis. Agric.*, 21, 1121–1155. <https://doi.org/10.1007/s11119-020-09711-9>.

Papageorgiou, E.I., Angelopoulou, K., Gemtos, T.A., Nanos, G.D. (2018). Development and evaluation of a fuzzy inference system and a neuro-fuzzy inference system for grading apple quality. *Applied Artificial Intelligence*, 32(3), 253-280. <https://doi.org/10.1080/08839514.2018.1448072>.

Zhang, J., Rao, Yu., Husband, K., Jiang, Z., Lee, S. (2021). Identification of cucumber leaf diseases using in-depth training and small sample size for the Agricultural Internet of Things. *International Journal of Distributed Sensor Networks*, 17, 1-13. <https://doi.org/10.1177/15501477211007407>.

Neupane, J., Guo, W. (2019). Agronomic bases and strategies of accurate water resources management: a review. *Agronomy*, 9, article number 87. <https://doi.org/10.3390/agronomy9020087>.

Zhao, W., Li, J., Yang, R., Li, Y. (2018). Determining placement criteria of moisture sensors through temporal stability analysis of soil water contents for a variable rate irrigation system. *Precision agriculture*, 19(4), 648-665. <https://doi.org/10.1007/s11119-017-9545-2>.

Lou, Yu-K., Sedler, E.J., Kemp, C.R. (2005). Economic justification of variable irrigation during corn cultivation on the south-eastern coastal plain. *Journal of Sustainable Agriculture*, 26, 69-81. <https://doi.org/10.3390/agronomy9020087>.

Mauget, S.A., Adhikari, P., Leiker, G., Baumhardt, R.L., Thorp, K.R., Ale, S. (2017). Modeling the impact of management and increase on cotton production in arid areas of West Texas. *Agricultural and Forest Meteorology*, 247, 385-398. <https://doi.org/10.1016/j.agrformet.2017.07.009>.

UDC 004.42:631.5:633

ANALYSIS OF THE METHODS AND SPECIFIC OF THE FIELD TRIAL DATA DIGITALIZATION AS THE BASIS OF FARMING MANAGEMENT

Volokha M., Dr. of Tech. Sciences, Prof.,

e-mail: volmp@i.ua, <https://orcid.org/0000-0002-0112-7324>

Yablonskyi P., candidate of technical Sciences, Assoc.,

e-mail: ypn@ukr.net, <https://orcid.org/0000-0002-1971-5140>

Lazarchuk M.,

e-mail: mizarchuk@ukr.net, <https://orcid.org/0000-0001-6192-6825>

Yu. Lazarchuk-Vorobiova,

e-mail: _jullazarchuk@gmail.com, <https://orcid.org/0000-0002-7866-3299>

Vorobyov O., s

e-mail: voroba.ua@gmail.com, <https://orcid.org/0000-0001-5314-1075>

Grubych M.,

e-mail: Mariya.grubich@gmail.com,

National Technical University of Ukraine «Ihor Sikorskyi Kyiv Polytechnic Institute»

Summary

The purpose of the study. *The topic of crop management, which is largely determined by modern digitalization processes, is relevant and is in the center of attention of both specialists and experts in the field of agriculture and in the field of computer technology, because the production of products of the agricultural sector plays a vital role in the world economy. Considering that traditional field data processing methods are unable to meet the ever-growing needs of agricultural producers at the new stage of agricultural development and are a serious obstacle to obtaining the necessary information, the purpose of the article is to conduct a critical review and analysis of publications on the digitization of field research databases in order to develop and adopt effective management decisions in crop production.*

Research methods. *Research was conducted using generally accepted scientific methods: abstract-logical; analysis and synthesis; induction and deduction; expert evaluations.*

Research results. *The conducted analysis made it possible to determine that the level of development of agricultural enterprises currently largely depends on modern digital technologies, the implementation of which involves a change in the general paradigm of production process management and allows commodity producers to act accordingly to increase production volumes. It has been proven that along with updating the material and technical component, the priority of production is the intellectualization of production and management activities based on digitization.*

Conclusion. *In order to benefit from the ever-increasing amount of data that comes from numerous sources of digital transformation, despite the fact that the vast majority of farmers and agricultural producers are not experts in this field and are unable to fully understand the basic laws of the algorithms being created, the scientific and methodological approach to increase the effectiveness of machine learning for automatic recognition of agricultural crops, detection of diseases and weeds, forecasting of yield and quality of the crop, management of water resources and soil can be useful for agricultural enterprises of many countries of the world.*

Keywords: *machine learning, precision farming, productivity, soil conditions, water resources.*