

# КІБЕРНЕТИКА та КОМП'ЮТЕРНІ ТЕХНОЛОГІЇ

УДК [502:519.7]+57.08:581.1

DOI:10.34229/2707-451X.22.2.10

Є.В. БАБЕНКО

## МЕТОДИЧНІ ОСНОВИ ПРОЕКТУВАННЯ ІНФОРМАЦІЙНОЇ СИСТЕМИ У РОСЛИННИЦТВІ

**Вступ.** Створення нових технологій для прецизійного землеробства призначене підвищити продуктивність, оперативність праці та поліпшити виробничі процеси. Інвестиції у сучасній технології – необхідність щоб відповідати різноманітним стандартам, продукція сучасних фермерських господарств має бути конкурентоспроможною на світовому ринку. Що у свою чергу дає можливість зменшити кількість працівників у сільському господарстві. Умови сьогодення вимагають від сільського господарства прогнозувати планету з населенням кількість якого постійно зростає, мінімізувати витрати та розробити технології, які не забруднюють планету.

Прецизійне землеробство насамперед побудоване на позиціонуванні за допомогою супутникової навігації, використанні сучасних продуктів для спостереження Землі, таких як карти, цифрові моделі та аналіз, а також цифрового зв'язку для передачі даних, наприклад, для відстеження худоби, вимірювання та контролю різних параметрів. Навігаційні системи, картографічні системи та метеостанції, а також управління операціями, комп'ютерне керування виробництвом кормових сумішей, сушіння зерна чи автоматичне керування споживанням на біогазових станціях, або моніторинг стану здоров'я худоби за допомогою спостереження за вагою в умовах реального часу, основний перелік технологій, які найчастіше використовуються у сільському господарстві.

Конфлікти та економічні потрясіння завжди відображаються на недоступності та дефіциті їжі у багатьох країнах світу. Крім того, світові запаси родючих земель обмежені геополітичними стратегіями гравців світової економіки. Нобелівську премію миру 2020 року присуджено Всесвітній продовольчій програмі (WFP – World Food Programme). За даними цієї програми, 811 мільйонів людей хронічно голодують, 283 мільйони перебувають у стані голоду або наближаються до голодомору. Приблизно ще 45 мільйонів у 43 країнах по всьому світу відчувають надзвичайний голод, іншими словами, голод вже чекає на них [1].

*Розглянуто основні підходи до побудови повномасштабного експерименту з точки зору методології планування, обробки даних, вибору моделі. Розкрито методичні основні, принципи та практична складова планування експерименту у рослинництві. Представлені власні результати розробки системи автоматичного прийняття рішень та базові математичні моделі, щодо побудови інформаційної системи.*

**Ключові слова:** бездротова сенсорна мережа, методологія, біосенсори, інформаційна система, математична модель, сільське господарство, фізіологія рослин.

© Є.В. Бабенко, 2022

Тому моделювання біологічних об'єктів, дослідження та проектування інтелектуальних систем для сільського господарства сьогодні викликають велику зацікавленість у вчених усього світу.

Побудова нових інформаційних технологій передбачає ряд попередніх етапів розробки, проектування, створення моделей досліджуваного явища чи об'єкта. До методів моделювання належать польове моделювання, фізичне моделювання, математичне моделювання. Проте моделювання біологічних систем – дуже складне завдання. Існують біологічні, фізико-хімічні, математичні моделі. В основному біологічні моделі відтворюють роботу окремих органів чи систем [2]. Моделі – це засоби, що використовують різні види впливу, генетичний апарат, мікробні впливи, токсини, хірургічні методи, гормони. Вони відтворюють фізичними чи хімічними засобами біологічні структури, функції, процеси, які на жаль, дають відповіді лише на конкретні питання і далекі від розуміння системи у цілому [3–5].

Особливий інтерес викликає вивчення процесу фотосинтезу як системи утворення речовини шляхом перетворення сонячної енергії, а також вивчення механізму, за допомогою якого рослини можуть отримувати енергію. Можна припустити, що створивши модель штучного фотосинтезу, в якому за допомогою основних процесів фотосинтезу можна створювати нові види палива, запобігати негативному впливу CO<sub>2</sub>, годувати планету. По суті, рослина з точки зору фізики процесу є акумулятором, на думку провідних вчених, які працюють над проблемами, пов'язаними зі створенням штучного фотосинтезу, можна створити новий тип акумулятора для супутникових систем [6, 7].

У попередній роботі нам вдалося створити промисловий зразок приладу «Флоратест», який працює на принципах ефекту Каутського [8]. У подальшій роботі ми створили сенсорну систему [9]. Нині ми працюємо над проектуванням інтелектуальної системи, практичними розробками баз знань інтелектуальної системи.

**Основи планування дослідження.** Основою вивчення рослинного організму є натурний експеримент. Натурні експерименти у рослинництві важко відтворюються і стикаються з низкою складнощів пов'язаних з точністю методів, які були використані для вивчення. Точність методів полягає у вірогідності та прецизійності результатів виміру. До поняття вірогідності належить відповідність результату натурного експерименту з арифметичною середньою великою кількістю результатів вимірювання та дійсним або прийнятим еталонним значенням. Щодо прецизійності, це ступінь відповідності між результатами експериментів. Необхідно враховувати прецизійність при проведенні натурного експерименту, передбачається, що дослідження, що проведені на однаковому матеріалі за незмінних умов, насправді може призводити до повторюваності результатів. Причиною є неминучі випадкові помилки, які є невідмінною складовою всіх вимірювальних підходів та процесів.

Чинники, які впливають на результати вимірювання, не можна ніколи повною мірою контролювати. При практичній інтерпретації вимірювальних даних завжди треба брати до уваги їхню варіабельність. В експерименті може статися так, що різниця між результатами експерименту і будь-яке окреме значення буде у тому ж діапазоні, що і неминучі випадкові помилки. В такому разі фактичне відхилення від заданого значення не може бути визначене. Аналогічно порівняння результатів дослідження рослин одного й того ж сорту може не вказати на принципову різницю врожайності, якщо результати можна віднести до змін, які є частиною процедури вимірювання. Багато інших факторів можуть сприяти мінливості результатів окремого методу вимірювання (навіть якщо вони враховують різницю між досліджуваними зразками, які вважають однаковими). До цих факторів можна віднести, наприклад, експериментатора, обладнання, калібрування обладнання, навколишні умови (температура, вологість, забрудненість повітря і т. д.), час, час який проходить між вимірюваннями. Значення окремої ознаки, отримане в результаті одного спостереження, називається спостережуваним значенням  $x_i$ . Результат тесту  $y_i = f(x_i)$  – це

значення характеризує, отримане при виконанні зазначеного методу тестування, яке може визначати фізичний розмір, кількість експериментів тощо.

Також існує варіабельність між вимірюваннями, виконаними різними операторами та/або з використанням різних пристроїв. Зазвичай це значення буде більше, ніж варіабельність між вимірюваннями, виконаними одним і тим самим оператором протягом короткого періоду часу з використанням одного і того ж пристрою. У варіабельності між вимірами, що повторюються, використовується термін *прецизійність*.

Для опису мінливості методів вимірювання у багатьох практичних випадках виявилось необхідним та корисним визначити дві умови *прецизійності*, які називають умовами повторюваності та умовами відтворюваності. В умовах повторюваності із зазначеними вище факторами (експериментатор, використане обладнання, калібрування обладнання, навколишні умови, час, що проходить між вимірами) вважаються постійними і тому не роблять внесок у мінливість, тоді як в умовах відтворюваності ці фактори змінюються і таким чином роблять внесок у мінливість результатів випробувань.

Таким чином, повторюваність та відтворюваність є крайніми випадками *прецизійності*, перша описує найменшу, а друга найбільшу мінливість результатів. Між цими крайнощами можливо представити й інші умови, коли один або декілька факторів можуть змінюватись і ці проміжні умови використовуються у окремо обмежених обставинах.

*Прецизійність* зазвичай виявляється у вигляді стандартних відхилень. Якщо можна розглянути справжнє значення вимірюваної властивості, то цікавить вірогідність методу виміру. Хоча справжнє значення не може бути відоме для деяких методів вимірювання, може бути прийнято еталонне значення вимірюваної властивості, наприклад, якщо доступні відповідні стандарти або якщо еталонне значення може бути визначено на основі іншого методу вимірювання або шляхом підготовки вже відомого зразка.

Вірогідність методу виміру може бути перевірена шляхом порівняння прийнятого еталонного значення з рівнем результатів, що забезпечуються обраним методом виміру. Вірогідність зазвичай виражається відхиленням, яке може статися, наприклад, у хімічному аналізі, якщо метод вимірювання дає збій або якщо один елемент повинен бути повністю виключений, або якщо визначення одного елемента впливає на присутність іншого. Термін *точність* деякий час застосовувався лише однієї складової, яку сьогодні називають *вірогідністю*. Для багатьох вона може означати відхилення результату від еталонного значення, викликаного як систематичними, так і випадковими впливами. У вимірах використовують базову статистичну модель. При оцінці *прецизійності* методу вимірювання доцільно припустити, що кожен результат випробувань – сума трьох складових для конкретного дослідного матеріалу:

$$y = m + B + e, \quad (1)$$

де  $m$  – загальне середнє,  $B$  – відхилення, що вноситься лабораторією за умов повторюваності,  $e$  – випадкова помилка, що виникає при кожному вимірі за умов повторюваності. Загальне середнє значення  $m$  – тестовий рівень, наприклад, зразки з хімічними речовинами різної чистоти або матеріалами різних типів, відповідатимуть різним рівням. У багатьох технічних ситуаціях тестовий рівень визначається виключно методом вимірювання, поняття незалежного істинного значення не може бути використане [10].

Загальне середнє значення  $m$  – це рівень тесту в експерименті *конкордантності*, який може не дорівнювати справжньому значенню  $\mu$ . Можливе відхилення обумовлено зміщенням методу  $\delta$ .

$$m = \mu + \delta. \quad (2)$$

При порівнянні результатів тесту зі значеннями, зазначеними у настанові, проводиться порівняння з істинним значенням  $\mu$ , яке може бути прийнятим за еталонне значення.

$$m = \frac{1}{n} \sum_{i=1}^n \frac{1}{n_i} \sum_{j=1}^{n_i} y_{ij}. \quad (3)$$

За постійних умов спостереження, зміщення  $B$  вважається постійним, тоді як зміщення змінюється при зміні оператора або вимірювальних методів та обладнання (зазвичай лабораторії). Зміщення  $B$  – сума випадкових і систематичних компонентів, а ступінь його мінливості відображається значенням дисперсії  $\sigma_L^2$ :

$$\text{var}(B) = \sigma_L^2. \quad (4)$$

Тоді загальна лабораторна похибка:

$$\Delta = \delta + B. \quad (5)$$

Випадкова похибка  $e$  виникає при кожному результаті випробування (загалом, внутрішньолабораторна мінливість залежить від умов роботи), ступінь її мінливості виражається значенням дисперсії  $\sigma_W^2$ :

$$\text{var}(e) = \sigma_W^2. \quad (6)$$

Різниця у мінливості випадкової похибки між лабораторіями невелика, і тому виправданим є визначення середньої мінливості методу вимірювання. Це значення називається дисперсією повторюваності  $\sigma_r^2$  і виконується:

$$\sigma_r^2 = \overline{\text{var}(e)} = \overline{\sigma_W^2}. \quad (7)$$

Ступінь узгодження результатів випробування відображається стандартним відхиленням повторюваності  $\sigma_r$  і стандартним відхиленням відтворюваності  $\sigma_R$ , для яких:

$$\begin{aligned} \sigma_r &= \sqrt{\text{var}(e)}, \\ \sigma_R &= \sqrt{\sigma_L^2 + \sigma_r^2}. \end{aligned} \quad (8)$$

Ступінь узгодження також можна виразити значенням оцінки стандартного відхилення  $S_r$  і коефіцієнта варіаційної повторюваності  $v_r$  (для одного методу, одного оператора в одній лабораторії) і відтворюваності  $S_R$  і  $v_R$  (між операторами і лабораторіями). Перевірка різниці між результатами випробувань базується на оцінці критичної різниці. Стандарт визначає межу повторюваності  $r$  (для одного методу, одного оператора в одній лабораторії) і межу відтворюваності  $R$  (між операторами та лабораторіями), як різницю між двома результатами тесту на рівні ймовірності 95%. Максимальна різниця між результатами двох тестів належить до межі повторюваності:

$$\begin{aligned} r &= f^* \sqrt{n^*} \sigma_r, \\ R &= f(n) * s_r. \end{aligned} \quad (9)$$

Якщо визначене абсолютне значення різниці між двома групами результатів або групою результатів і справжньою величиною (діапазоном) більше за критичну різницю, результати тесту є підозрілими. У такій ситуації стандарт передбачає проведення подальшого випробування того ж матеріалу в умовах повторюваності. Розширення кількості вимірювань допускається лише один раз, при цьому схема враховує витрати на отримання результатів випробувань. У таблиці приведені варіанти тестів, з яких можливі два наступні для недорогих, та для дорогих тестів дозволяється збільшення на один результат.

ТАБЛИЦЯ. Дослідження у залежності від витрат на проведення тестів

№ 1	№ 2
2 результати виміру	$n \geq 5$ результатів виміру
Рішення	Рішення
Наступні 2 результати вимірювання	Наступні $m$ ( $n/3 < m < n/2$ ) результатів вимірювання
Остаточний записаний результат, $2n$ , медіана, $S_R$	Остаточний записаний результат, $n + m$ , медіана, $S_R$

При оцінці прецизійності дослідження треба починати з його планування та взяти до уваги відповіді на наступні питання:

- чи є підходящий стандарт для способу виміру?
- скільки лабораторій/груп дослідників потрібно залучити до спільної роботи над експериментом?
  - як залучити лабораторії та яким вимогам вони мають відповідати?
  - які діапазони рівнів зустрічаються на практиці?
  - скільки рівнів необхідно використовувати в експерименті?
  - які відповідні матеріали мають ці рівні репрезентувати та як їх треба підготувати?
  - яку кількість експериментів треба провести?
  - який графік встановити для проведення експерименту?
  - чи підходить базова статистична модель чи треба враховувати її модифікації?
  - особливу увагу слід приділити тому, чи буде передбачуваний матеріал знаходитися в тому самому фізичному та біохімічному стані при вимірюванні у всіх лабораторіях?

Наступним етапом планування є те, що запропонований метод вимірів повинен бути стандартизований. Такий метод має бути надійним, це означає, що невеликі відхилення у процедурі вимірювань не повинні призводити до несподівано великих змін у результатах. Якщо такі зміни можуть вплинути на результат, така методика має містити необхідні запобігання або попередження. При розробці стандарту для окремого метода вимірювань доцільно зосередити всі зусилля на усунення або мінімізацію відхилень. Аналогічні процедури можуть використовуватися для оцінки вірогідності та прецизійності, як у випадках встановлених методів вимірювання, так і у разі нових стандартизованих методів вимірювань. У разі нових стандартів результати слід розглядати, як попередні оцінки, оскільки вірогідність та прецизійність змінюватимуться у міру накопичення лабораторіями досвіду. Документ, що описує методику, має бути однозначним та повним. Усі основні операції, що стосуються навколишнього середовища, в якому проводиться процедура вимірювань, реагентів, апаратури, попереднього тестування обладнання та підготовки досліджуваного зразка, повинні бути включені до методики вимірювання. Або можливо навести посилання на інші описані процедури подібних вимірювань, які доступні операторам. Слід зазначити метод розрахунку та представлення результатів випробувань, включаючи кількість

необхідних повторень, достатніх для отримання вірогідних результатів. Виробнича специфікація може вимагати проведення повторних вимірів за умов повторюваності. У цих випадках можна застосувати стандартне відхилення для перевірки прийнятності результатів вимірювань та прийняття рішення про те, які дії необхідно вжити, якщо результати є невірогідними. Якщо результати вимірювання при стандартному застосуванні та застосуванні в умовах замовника відрізняються, при оцінці одного і того ж матеріалу, можна використовувати стандартне відхилення повторюваності та відтворюваності, щоб зрозуміти, чи знаходяться результати в діапазоні вірогідності, що передбачено методом вимірювання. Для вимірювання тої самої властивості можуть бути доступні два методи вимірювання, один з яких має бути простіше і дешевше іншого, але менш підходящий для цілей дослідження. Значення вірогідності та точності результатів можуть бути виправдані використанням менш дорогого методу для певної групи матеріалів.

**Матеріали дослідження, методи та моделі.** В основу інтелектуальної системи покладена задача отримання нових знань, де завдяки введенню нових вузлів є можливим розширення функціональних можливостей системи, а саме отримувати знання, що утворюються автоматично без втручання людини.

Вирішення поставленої задачі досягається тим, що система містить бездротову сенсорну мережу, онтолого-керований вузол [11], базу даних, базу знань, керуючий модуль, персональний комп'ютер та людино-машинний інтерфейс.

Введення у систему нових вузлів та зв'язків дозволяє розширити функціональні можливості системи, зокрема, отримати нові знання автоматично без втручання людини.

На рис. 1 показано структурна схема запропонованої системи. Яка містить бездротову сенсорну мережу 1, вихід якої через онтологічно керований вузол 2 з'єднаний із входом глобальної бази даних 3, вихід якої через онтологічно керований вузол 2 з'єднаний із входом бази знань 5, вузол 2 підключений до входу пояснювального модуля 6, вихід модуля керування якого є цілим пристроєм до комп'ютера та людино-машинного інтерфейсу 4, що дозволяє приймати 7 управлінських рішень.

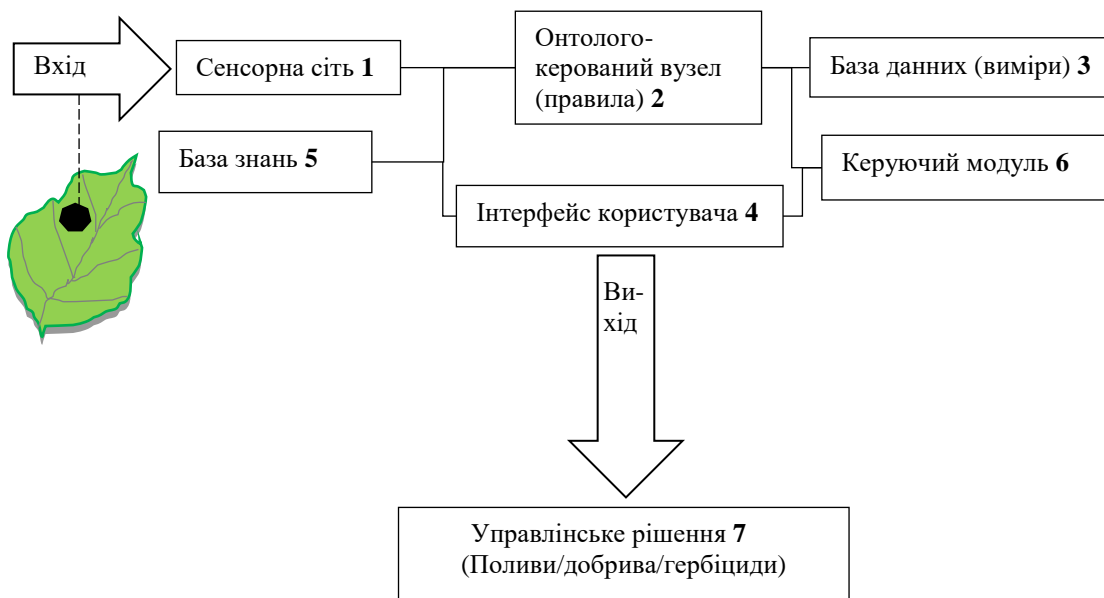


РИС. 1. Структурна схема інтелектуальної системи

У своїй роботі ми використовували методи моделювання даних, отриманих за допомогою натурального експерименту. Статистик G.E.P. Voxe сказав: «Усі моделі неправильні. Але деякі з них корисні», принцип пояснення будь-якого явища за бритвою Оккама, має бути максимально простим. Кожна модель є гіршою, ніж найкраще наближення до реальності [12]. Статистичне моделювання, обробка даних та візуальне представлення виконувалось у середовищі R (R-Studio).

Дослідним шляхом було обрано узагальнені лінійні моделі GLM (Generalised Linear Models), що складаються з трьох частин, лінійного предиктора, відповідних даних (експоненційний клас), лінія (трансформаційна функція, часто нелінійна). На початку розглянемо узагальнені лінійні моделі та моделі регресійного аналізу для моделювання статистично залежних, корельованих даних. Модель регресії має систематичну та випадкову складову. Наприклад, якщо нас цікавить відносини між вагою рослини ( $y$ ) та віком ( $x$ ) рослини, яку можна описати прямою лінією з двома параметрами ( $\alpha, \beta$ ), таким чином можливо до систематичної суми ( $\alpha + \beta x$ ) додати випадкову складову, наприклад, позначимо її буквою  $\varepsilon$  (теоретична випадкова мінливість). Таким чином, отримуємо:

$$y_i = \alpha + \beta x_i + \varepsilon_i. \quad (10)$$

Модель можливо легко розширити до вигляду загальної лінійної моделі LM (General Linear Model):

$$y = \alpha + \beta_1 x_1 + \beta_2 x_2 + \dots + \beta_k x_k + \varepsilon, \quad (11)$$

$\varepsilon \sim N(0, \sigma^2)$ , незалежно для різних вимірів.

Можемо також за допомогою лінійного предиктора моделювати невідповідність у вимірах ( $y$ ), а в перетворених середніх значеннях  $f(\mu)$ :

$$f(\mu) = \alpha + \beta_1 x_1 + \beta_2 x_2 + \dots + \beta_k x_k. \quad (12)$$

Для відображення результатів у графічній формі існують шість основних видів графів: Residuals versus fitted/predicted, Normal Q-Q, Scale-Location, Cook's distance, Residuals versus leverage, Cook's distance versus leverage.

Виходячи із специфіки досліджень при повторних спостереженнях (clustered data, panel data, longitudinal data, repeated measures, при повторних спостереженнях у часі) крім суб'єктів, у тому сенсі говорять про неспоріднені групи (clusters) спостережень та корельованих спостережень усередині групи. Взаємозалежність між спостереженнями у межах однієї і тієї ж області або предмета може виникнути, тому що спостереження просторово близькі. Через близькість повторних спостережень всередині суб'єкта (і дуже схожих умов середовища) можна очікувати, що субодиноці всередині одного кластера якимось взаємопов'язані, взаємодіють один з одним або якимось схожі. Часто, але не завжди наростання залежності по відстані монотонно, чим є ближче субодиноці одна до одної, то більше схожість і навпаки, чим вони далі одна від одної, тим менше залежні їх спостереження. Також велике значення має масштаб експерименту, з якого отримано дані. Дуже важливим аспектом при виборі моделі для подальшого аналізу є вибір, що є предметом дослідження. Ця залежність (кореляція), яка вводить наша статистична модель та її оцінка необхідна тому, що конкретним чином враховує кількість інформації, що міститься в експериментальних даних. Корельовані дані можна також використовувати, щоб оцінити різні характеристики, що виникають внаслідок залежності між спостереженнями. Створення спеціальної

експериментальної схеми спрощення оцінки чи порівняння контрольованих параметрів. Так, наприклад, варіабельність у висоті рослини можна уявити, як різниця між окремими вегетаційними камерами, як різницю між окремими горщиками всередині однієї вегетаційної камери, як різницю між рослинами одного горщика у межах одного вегетаційного боксу. Далі для оцінки мінливості між представниками одного сорту окремих рослин, між показниками одних і тих самих рослин, мінливість кінцевої маси відповідно до лінійної тенденції зростання однієї і тієї ж рослини.

Повністю рандомізований дизайн CRD (Completely Randomised Design) означає, що якщо ми порівнюємо кілька експериментальних варіантів, ми випадково відносимо їх до обраних експериментальних одиниць, наприклад, шляхом жеребкування або генерацією випадкових чисел.

Як приклад, розглянемо просту модель у якій залежна змінна  $y$  виходить із нормального розподілу і залежить від однієї пояснюючої змінної  $x$ . Для  $i$ -того виміру,  $i = 1, \dots, n$ , маємо:

$$y_i = \alpha + \beta x_i + \varepsilon_i, \quad (13)$$

де  $\varepsilon \sim N(0, \sigma^2)$ ,  $cor(\varepsilon_i, \varepsilon_{i'}) = 0$  для  $i \neq i'$ .

Припускаємо, що помилки мають нормальний розподіл, гомоскедастичні (мають однакову дисперсію) та незалежні одна від одної. Однак для обробки комплексних даних, які часто спостерігаються зі значним шумом підходять дві групи статистичних методів, параметрична і непараметрична лінійна регресія. Для конкретної моделі, з якою працюємо, треба розуміти математичний сенс окремих параметрів [13]. Функцію починаємо  $f$ , а в дужках зазвичай вводять імена параметрів, що змінюються,  $f(x, \theta)$ , де  $x$  вектор описує відомі мінливі дані,  $\theta$  невідомий параметр, який буде отримано з даних. Зазвичай параметрів більше ніж один, наприклад,  $k$ , а тоді

$\theta$  беремо, як векторний параметр, так  $\theta = \begin{pmatrix} \theta_1 \\ \vdots \\ \theta_k \end{pmatrix}$ .

Існує нескінченно багато функцій залежності, але часто використовуються всього декілька – статична функція, раціональні, експоненційні, логарифмічні та тригонометричний клас функцій.

Для наших цілей (оцінки індукції флуоресценції хлорофілу) використовували множинну непараметричну регресію. Модель виду:

$$\mu_i = \alpha + f_1(temp_i, rh_i), \quad (14)$$

$i_i \sim N(\mu, \sigma^2)$ ,  $cor(\varepsilon_i, \varepsilon_{i'}) = 0$ , де  $\alpha$  єдиний параметричний член, а  $f_1(temp_i, rh_i)$  функція двох змінних.

На рис. 2 показано залежність швидкості внутрішнього росту індукції флуоресценції хлорофілу (відносні одиниці) від дії гербіциду Гліфосат, температури (°C) та вологості (%). Як видно з оброблених експериментальних даних, найбільш ефективна дія при вологості повітря 75,7 % і температурі навколишнього середовища 25–30 °C. Нам вдалося експериментально відокремити математичну модель, яка успішно опрацювала дані, що отримані дослідним шляхом. На цьому етапі експериментальних досліджень використаний нами підхід до вибору моделі виявився працездатним. Ми вирощували рослини, зберігаючи необроблені параметри даних, а потім застосовували обробку за допомогою математичних моделей. Як фактор впливу, який ми вивчали, обрали дію гербіциду на рослини виду *Datura stramonium* L.



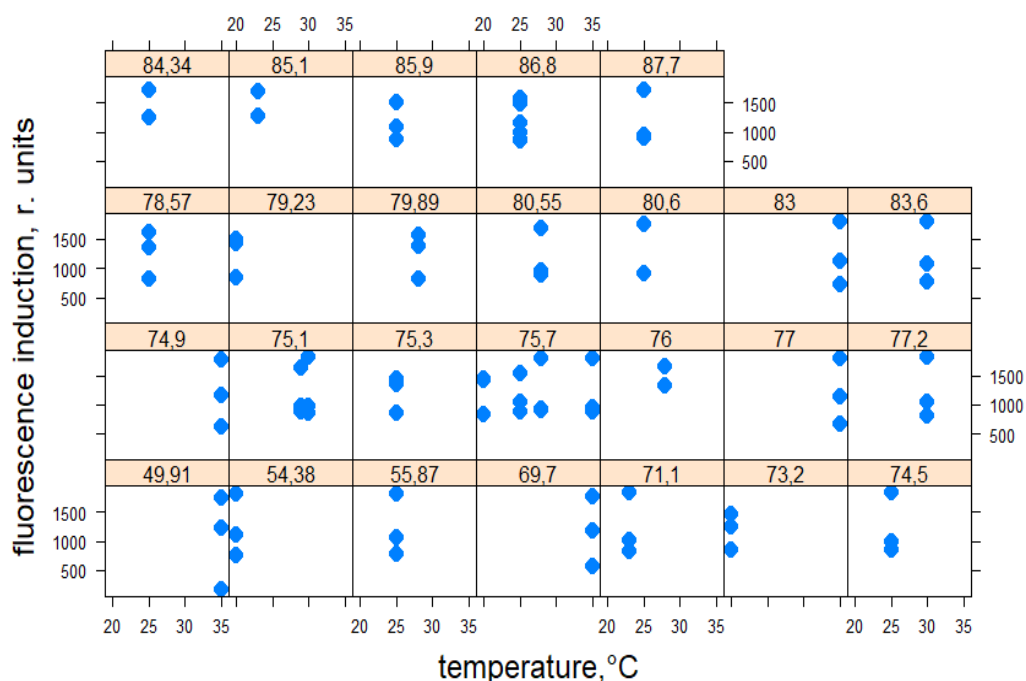


РИС. 2. Дії гербіциду на рослини в залежності від температури та вологості повітря

Наступний крок, потрібно заповнити базу даних різними вимірами необроблених даних, щоб наповнити онтолого-керований вузел.

Наукова новизна полягає у створенні методичної основи для проведення експериментальних робіт зі збору даних з бездротової сенсорної мережі та подальшого моделювання інформаційної системи, та проектування систем у рослинництві за запропонованою на рис. 1 функціональною схемою. Успішне впровадження зазначених у дослідженні підходів, дозволить нам перейти на принципово нову основу рослинництва, яка базується насамперед на потребах рослин та підвищенні врожайності. Наступними етапами наших досліджень є впровадження результатів у партнерських дослідженнях. Запланована співпраця з Науково-дослідним інститутом садівництва та харчових технологій, Кишинів, (Молдова). Заплановано експериментальне використання інформаційної системи для контрольованого зрошення виноградників. У наших попередніх роботах ми мали досвід використання розумних біосенсорів для точного землеробства [14] та для регулювання водного режиму [15]. З боку моделі це буде введення додаткових параметрів за допомогою методу бета-сепарації.

**Висновки.** Вивчення рослинних об'єктів зустрічається з рядом труднощів. У зв'язку з тим, що рослинний об'єкт є не лише сукупністю окремих систем всередині організму, а й піддається постійному впливу зовнішніх факторів оточуючого середовища, клімату та ґрунту, що необхідно враховувати при створенні нових інформаційних систем, ціллю яких є підвищення врожайності. Основою вивчення рослинного організму є натурний експеримент. Польові дослідження у рослинництві важко відтворити і стикаються з низкою труднощів, пов'язаних з точністю використовуваних методів дослідження, що полягають у вірогідності та точності результатів вимірювань.

Наразі доступно багато джерел даних, наприклад, сенсорні мережі, Інтернет речей, технологія блокчейн. Ми розглянули алгоритми обробки даних. Розглянута система алгоритмів здатна враховувати мінливість змін у багатофакторному середовищі.

Важлива частина цієї роботи – вивчення ефекту індукції флуоресценції хлорофілу та вивчення методології досліджень з фотосинтезу. Цей ефект дуже чутливий до багатьох змін у рослині. Це перевага методу та вимога до методології дослідження чи специфіки її проведення. Особливий інтерес представляє склад ґрунту та вплив характеристик ґрунту на ріст і розвиток рослин, для створення профільних ґрунтових карт. Вперше для дослідження використано багаторівневу інформаційну систему контролю за станом та потребами рослин, яка містить бездротову сенсорну мережу, онтологічно керований вузол, глобальну базу даних, базу знань, пояснювальний модуль, модуль керування, комп'ютер та людино-машинний інтерфейс, що дозволяє приймати управлінські рішення.

Представлено схему інформаційної системи, де модулем управління у режимі реального часу, виступає параметр індукції флуоресценції хлорофілу. Це перспективний шлях коригування режимів зрошення, моніторингу стану рослин та догляду за багаторічними насадженнями у сільському господарстві, зокрема, рослинництві.

#### Список літератури

1. [World Food Programme \(WFP\) – Facts – 2020 - NobelPrize.org](https://www.nobelprize.org/awards-and-prizes/2020/world-food-programme) (звернення 25.05.2022)
2. Dagan F.D. Biomedical Information Technology. Academic Press, 2020. 798 p. <https://doi.org/10.1016/C2017-0-03305-0>
3. Gisriel C., Sarrou I., Ferlez B., Kevin E. Redding and Raimund Fromme Structure of a symmetric photosynthetic reaction center–photosystem. SCIENCE, 27 Jul 2017. 357 (6355). P. 1021–1025. <https://doi.org/10.1126/science.aan5611>
4. Stowell M.H.B., McPhillips T.M., Rees D.C., Soltis S.M., Abresh E., Feher G. Light-Induced Structural Changes in Photosynthetic Reaction Center: Implications for Mechanism of Electron-Proton Transfer. SCIENCE, 2 May 1997. 276 (5313). P. 812–816. <https://doi.org/10.1126/science.276.5313.812>
5. Kaufmann M. Green Information Technology. 2015. 348 p. <https://doi.org/10.1016/C2014-0-00029-9>
6. Barber J. Photosystem II: the engine of life. Quarterly Reviews Of Biophysics. 2003. 36 (1). P. 71–89. <https://doi.org/10.1017/S0033583502003839G>
7. García de Arquer F. P., Talapin D.V., Klimov V.I., Arakawa Y., Bayer M., Sargent E.H. Semiconductor quantum dots: Technological progress and future challenges. SCIENCE, 6 Aug. 2021. 373 (6555). P. 1–14. <https://doi.org/10.1126/science.aaz8541>
8. Romanov V., Artemenko D., Brayko Yu., Galelyuka I., Imamutdinova R., Kytayev O., Palagin O., Sarakhan Ye., Starodub M., Fedak V. Portable Biosensor: from Idea to Market. *International Journal "Information Theories & Applications"*. 2012. 19 (2). P. 126–131. <http://www.foibg.com/ijita/vol19/ijita19-2-p04.pdf>
9. Romanov V., Galelyuka I., Sarakhan Ye. Wireless sensor networks in agriculture. The Seventh IEEE International Conference on Intelligent Computing and Information System, ICICIS 2015. Cairo, Egypt, 2015, 12–14 December. P. 79–83. <https://doi.org/10.1109/IntelCIS.2015.7397200>
10. ISO 5725-2:1994 Accuracy (trueness and precision) of measurement methods and results – Part 2: Basic method for the determination of repeatability and reproducibility of a standard measurement method. 42 p. <https://www.iso.org/standard/11834.html>
11. Сарахан Є.В., Палагін О.В., Романов В.О. Глобальна мережа для автоматичного отримання знань. Пат. UA.: 89025, бюл. №7/2014. <https://uapatents.com/5-89025-globalna-merezha-dlya-avtomatichnogo-otrimannya-znan.html>
12. Pekar S., Brabec M. Modern Analysis of Biological Data. 1. Generalized linear models in R 2<sup>nd</sup> edition. Masaryk University Press, Brno. 2020. 264 p.
13. Pekar S., Brabec M. Modern Analysis of Biological Data. 3. Non-Linear Models in R. Masaryk University Press, Brno. 2019. 218 p.
14. Romanov V., Artemenko D., Galelyuka I., Palagin O., Sarakhan Ye. Remote smart biosensors for precision farming and environment protection. *International Journal. Information Theories & Applications*. 2013. 20 (2). P. 174–179. <http://www.foibg.com/ijita/vol20/ijita20-02-p10.pdf>
15. Зеленьянська Н.М., Сарахан Є.В., Буркан Н.В., Тулинова Н.В. Флуоресценція хлорофілу та водний режим листків саджанців винограду. *Вісник аграрної науки: український науково-теоретичний журнал*. 2009. Том. 9. С. 25–27. <https://agris.fao.org/agris-search/search.do?recordID=UA2009000086>

Одержано 19.08.2022

**Бабенко Євгенія Володимирівна,**

кандидат сільськогосподарських наук, старший науковий співробітник  
Інституту кібернетики імені В.М. Глушкова НАН України, Київ.

<https://orcid.org/0000-0002-0983-9713>

UDC [502:519.7]+57.08:518.1

**Yevheniia Babenko**

## **Methodological Fundamentals of Information System Design in Crop Production**

*V.M. Glushkov Institute of Cybernetics of the NAS of Ukraine, Kyiv*

*Correspondence: [sarakan2006@ukr.net](mailto:sarakan2006@ukr.net)*

**Introduction.** The creation of new technologies for precision agriculture is intended to increase productivity, labor efficiency and improve production processes. According to the World Food Program (WFP), 811 million people are chronically hungry, 283 million are in a state of starvation or close to starvation. An estimated 45 million more in 43 countries around the world are on the brink of starvation. Today's conditions require agriculture to feed a planet with an ever-growing population, minimize costs, and develop technologies that do not pollute the planet. Therefore, modeling of biological objects, research and design of intelligent systems for agriculture are of great interest to scientists around the world today.

**The purpose** of the paper is development of the main approaches to building a full-scale experiment from the point of view of planning methodology, data processing, and model selection. The methodical basics, principles and practical component of planning an experiment in crop production are disclosed. The results of the development of the automatic decision-making system and the basic mathematical models for the construction of the information system are presented.

**Results.** The basis of the study of the plant organism is a natural experiment. Field experiments in crop production are difficult to reproduce and face a number of difficulties related to the accuracy of the research methods used, which consist in the reliability and accuracy of the measurement results.

For the first time, a multi-level information system for monitoring the condition and needs of plants, which contains a wireless sensor network, an ontologically controlled node, a global database, a knowledge base, an explanatory module, a control module, a computer, and a human-machine interface, which allows taking management decisions, was used for the research. Our algorithm system is able to take into account the variability of changes in a multivariate environment. An information system where the chlorophyll fluorescence induction parameter, induction, measured in real time, acts as a control module. This is a promising way of adjusting irrigation regimes, monitoring the condition of plants and caring for perennial plantations.

**Conclusions.** An important part of this work is the study of the effect of induction of chlorophyll fluorescence and the study of the methodology of research on photosynthesis. This effect is very sensitive to many changes in the plant. This is an advantage of the method and is a requirement for the research methodology or the specifics of its conduct.

The study of plant objects faces a number of difficulties. Of special interest is the composition of the soil and the influence of soil characteristics on the growth and development of plants, for the creation of profile soil maps. Due to the fact that the plant object is not only a collection of individual systems inside the organism, but is also exposed to the constant influence of external factors of the environment, climate and soil, which must be taken into account when creating new information systems, the purpose of which is to increase productivity.

**Keywords:** wireless sensor network, methodology, biosensors, information system, mathematical model, agriculture, plant physiology.